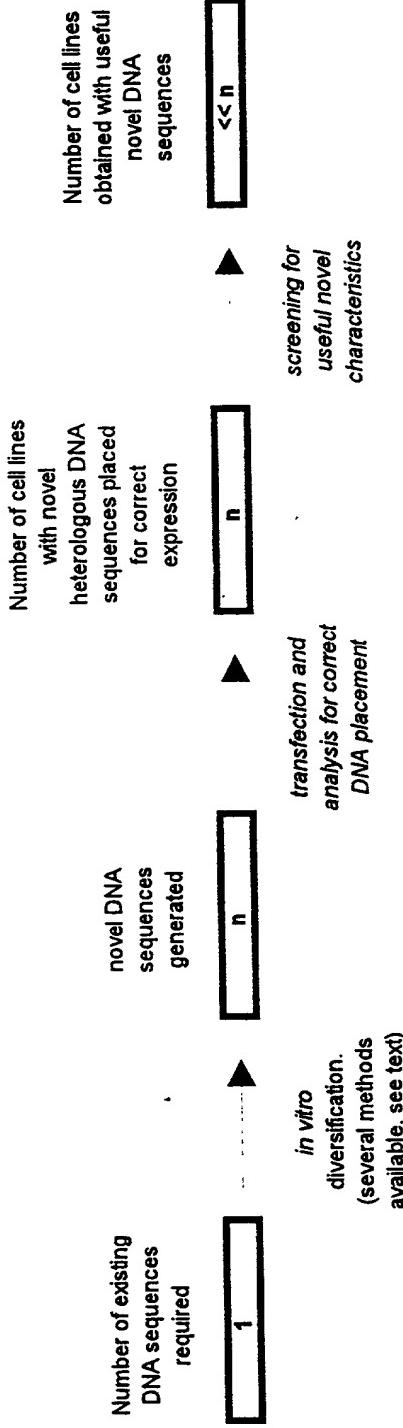


**Figure 1** Methods for the diversification of DNA sequences and testing for superior variants

11 According protocols: Number of transfusions needed to generate 1024 new variants: 1024

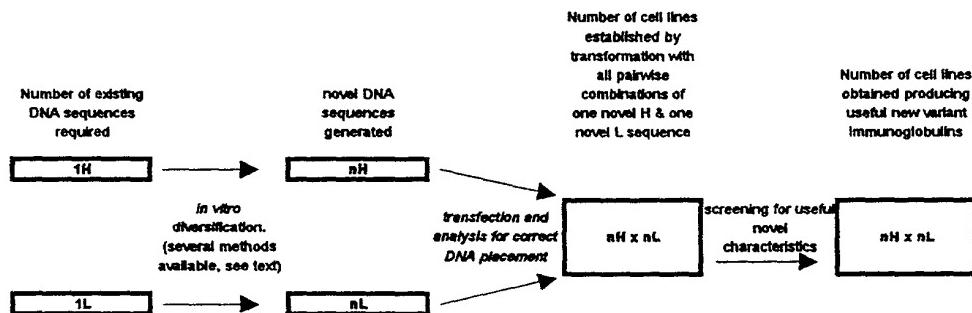


1B a protocol enabled by the present invention: Number of transsections needed to generate 1024 new variants: 2

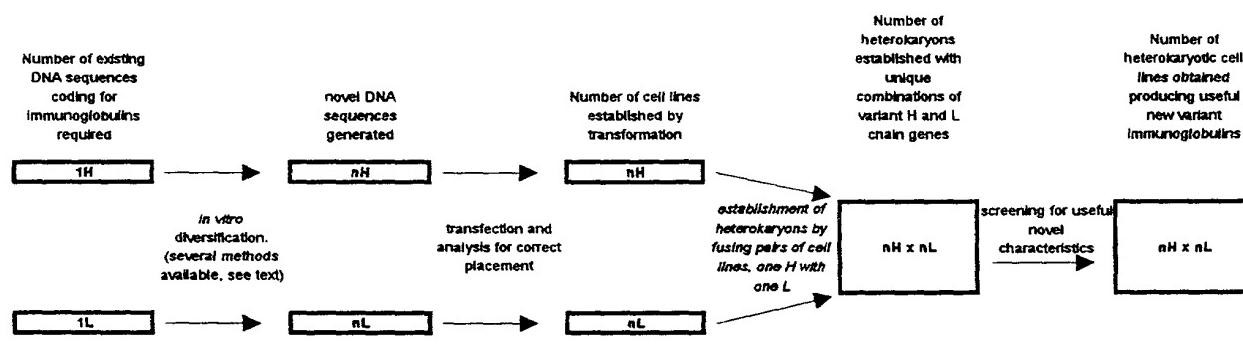
**Figure 2 Methods for the diversification of DNA sequences coding subunits of heteropolymeric proteins and testing for superior variants.**

The example given for immunoglobulins is for illustrative purposes only and is not intended to limit application of the present invention to this specific heteromeric protein. H = heavy chain genes, L = light chain genes

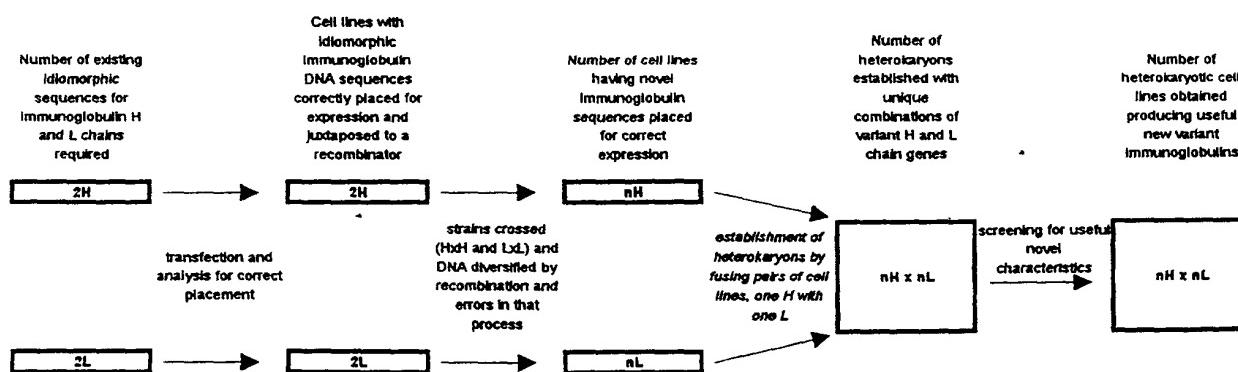
**Figure 2A existing protocol: Number of transfections needed to generate 1024 new combinations: 2048**



**Figure 2B existing protocol using the heterokaryon technology of US Patent Serial No. 5,643,745**  
**Number of transfections needed to generate 1024 new combinations: 64**

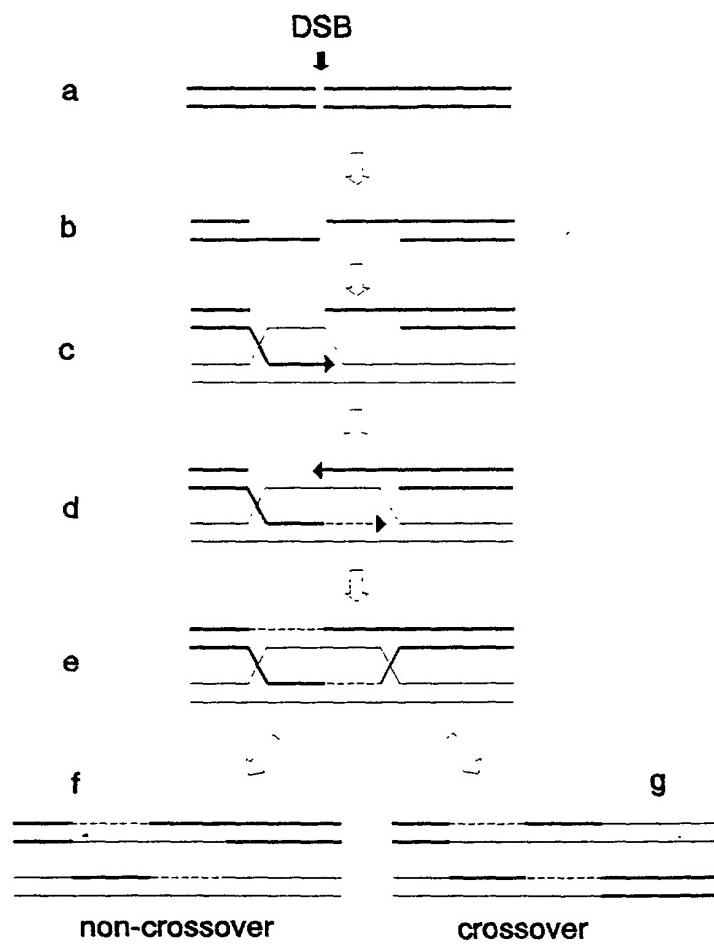


**Figure 2C a protocol enabled by the present invention**  
**Number of transfections needed to generate 1024 new combinations: 4**

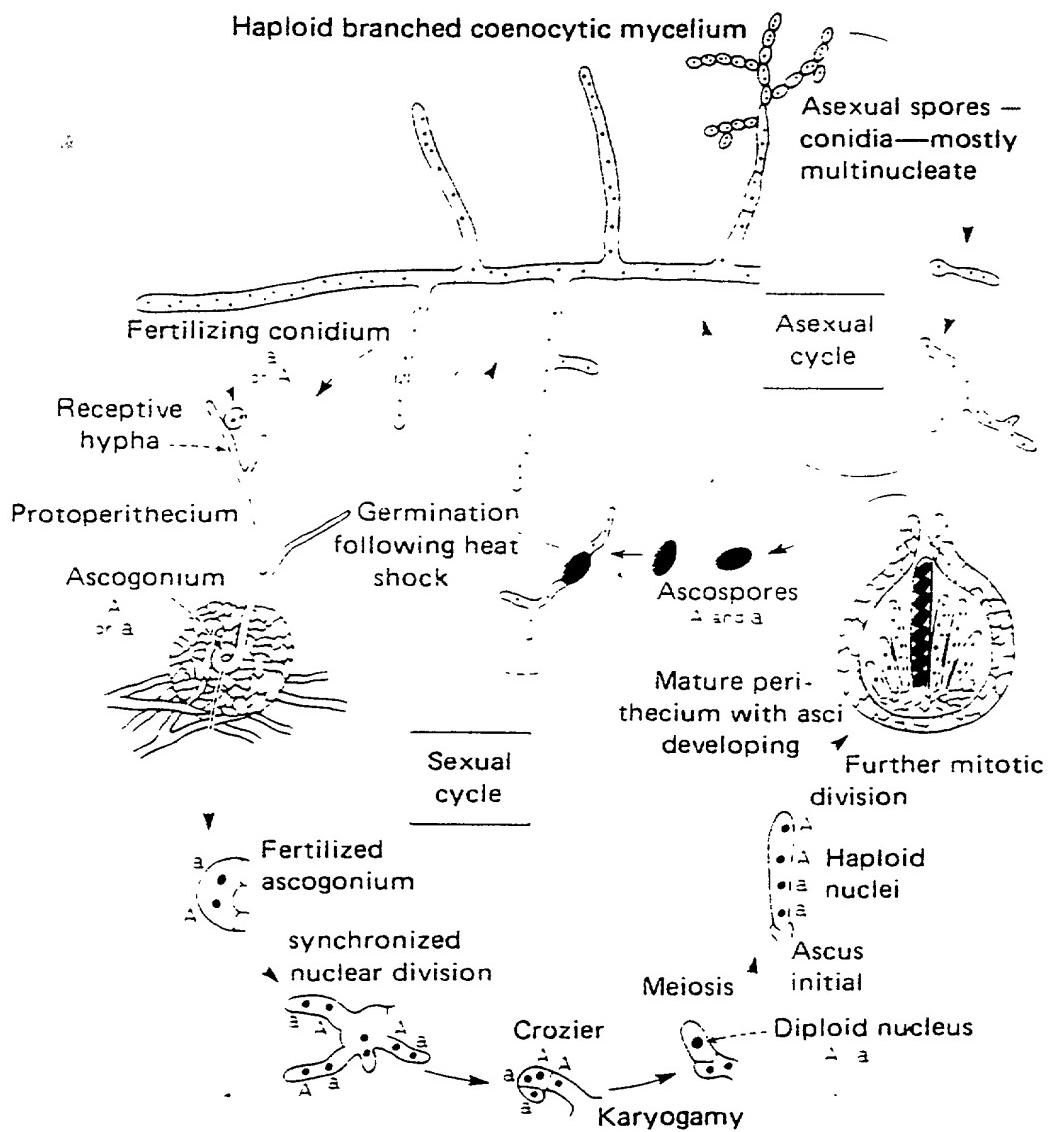


**Figure 3** The modified double strand break repair model for meiotic recombination. After H Sun *et al* Cell 64: 1155-1161, 1991

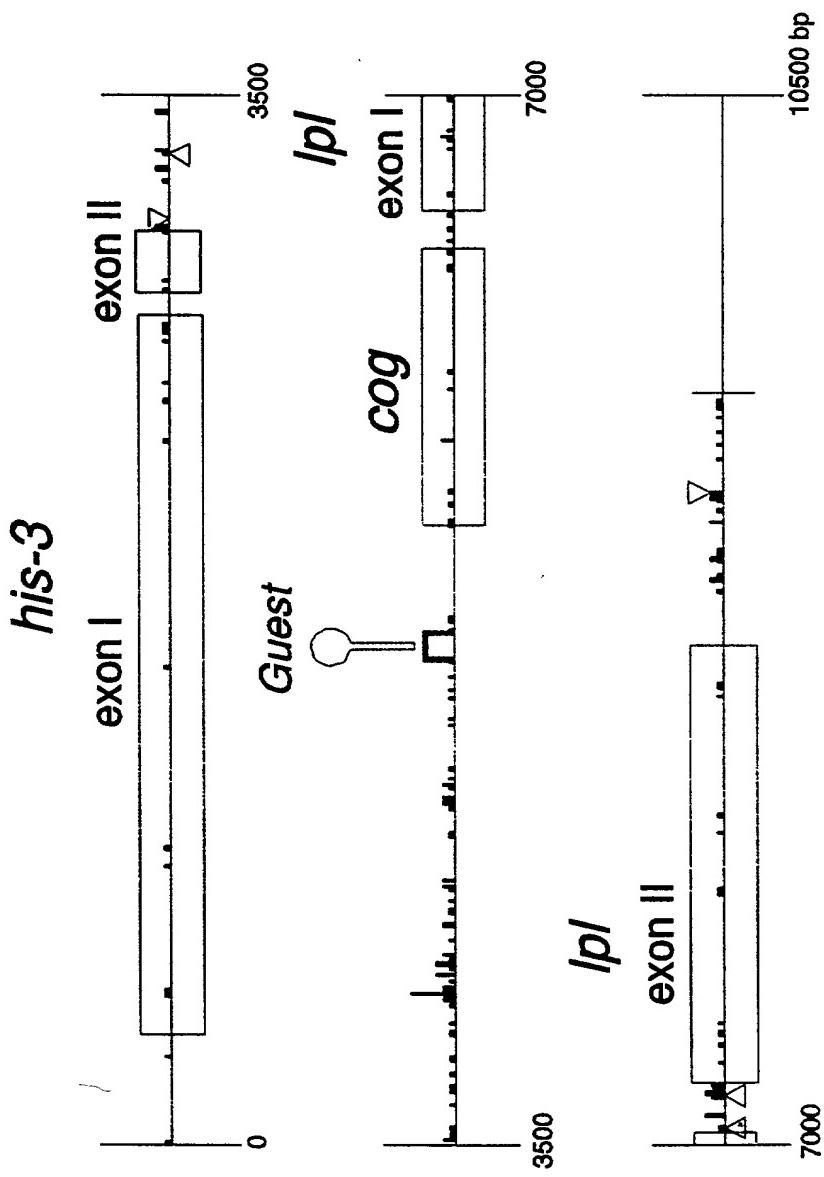
(a) A double strand break (DSB) is made in one DNA duplex. (b) A long 3' overhanging single strand tail is generated either side of the break by resection. (c) One 3' end invades a homologous duplex forming a D loop. (d) the D loop is enlarged by repair synthesis and anneals to the second 3' end. (e) Repair synthesis occurs at the second 3' end and two intermolecular junctions (Holliday junctions) are formed. Resolution of the junctions by cutting inner and outer strands can give rise to non-crossover (f) and crossover (g) chromosomes. If there are base mismatches in the heteroduplex regions (duplex molecules with thick and thin lines) there will be gene conversion. If mismatch repair does not occur there will be post meiotic segregation of new sequence combinations.



**Figure 4** Life cycle of *Neurospora crassa* after JRS Fincham (Genetics, Wright 1983). Microconidia having one nucleus are not shown but can be generated as described in the text. Perithecia and protoperithecia are shown in section.



**Figure 5** Map of the *his-3*, *cog*, *lpl* region of Linkage Group I of *Neurospora crassa*. Vertical bars, triangles and hairpins show the location of sequence differences that distinguish the St Lawrence and Lindegren wild type strains. The corresponding full DNA sequences are given in figure 7 and figure 8. Vertical slashes indicate one to seven base substitutions per 10 base pairs. Triangles indicate short sequence insertions and the hairpin a 101 base pair inverted repeat transposon fragment present in St Lawrence.

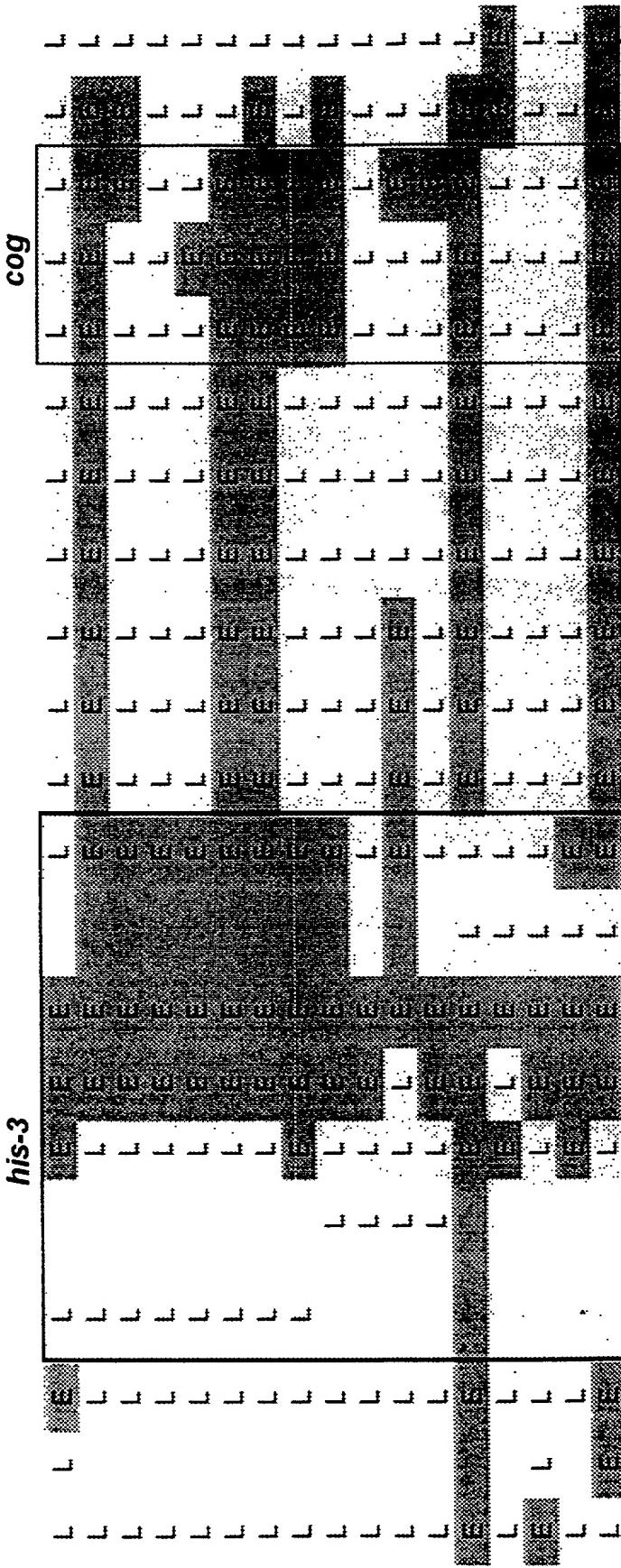


**Figure 6.** Discontinuity in the parental origin of DNA sequences in progeny from crosses between pairs of *his-3* alleles. In most cases this reflects discontinuity of conversion tracts, in some cases crossovers near the ends of conversion tracts.

Markers are specific DNA sequence differences that distinguish the parents. These were all E (Emerson wild type origin) or all L (Lindegren wild type origin) in the parental strains. Recombinants carry both E and L markers.

Marker position is given in base pairs from the first base of the *his-3* gene. Each line of the table shows the parental origin of the markers inherited by one of the progeny.

marker	P	H	P1	K1201	K504	L3	R1	K26	K874	R4	C4	C5	C6	C6/7	C7	C8	C1	C2	C3	C9	D
location	~600	-384	115	179	563	1232	1502	1717	2318	3436	3705	4000	4304	4667	4821	5232	5495	6153	6507		



**Figure 7** Nucleotide sequence of the *his-3 cog<sup>L</sup> lpl* region of linkage group I in the Lindegren wild type strain of *Neurospora crassa*. This differs from that in the StLawrence strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the high frequency recombinator *cog<sup>L</sup>* which is active providing the cross in which meiosis occurs is homozygous *rec-2*.

1	GATCGCAACT	GGAGATCACT	CGCACCGTGC	CGCAGAACAA	GGCGGACGAG	CCTCAGGGCA
61	GTTCAGCCTG	CCGTAACAGC	ACAGACCATA	GCTTATTTC	ACCTGGGCGG	GCGGGCGACG
121	GCAGCACTGA	CATCGGAAG	GCGGCATCAA	GCAACCCCTC	TGTTGCTTGC	CAGCTGCCGG
181	CCAACGTCAG	CGGTACAAGG	AGAAATCTGG	AAGGAAAGAC	TTCTGGCACC	GACAGGATGG
241	CACGGGGAA	AAGTCCC	TGCATGAGAT	GAGGGGCATT	TGCAATTGCCT	CCCGTCACAC
301	TGCCCGCAA	CCCCAACCCC	ACCATAGCGT	CTGTCGATAC	ATGGAGCGCG	AAGTCGAGAA
361	ACCTGTAATT	CCTGGTAACT	TTCAAGGTACA	CAGTACGTAC	TGATCCTGGT	ATCAAACCTT
421	GCCTGCCAG	TTTCGACGG	AAAGAGGTGT	GAATTGTGAA	AGAGTCATAC	CAAATCACCC
481	GATTTTCATA	AAGCCCAGT	CTTTTCTGTA	CATAAGCGAC	ACTCGAAGCG	GGCCTCATCT
541	TCATAGCCTG	ATAGCTTGT	ATACTCCATC	CTCGTATCTC	ACTTGACCTT	GAGTTCAACC
601	CCACGTCAGA	CTTCACCCGA	CACATCGACG	GATTGGGAA	CAGCACAATA	CCTGAAAAGC
661	GAGAAAACCA	AACAGAGGAA	AACACCATGG	AGACAACACT	TCCCCTCCCC	TTCCCTCGTCG
721	GTGTCACTG	TCCTCCCGGA	CTGAATGACA	TCAAGGAGGG	CCTCAGCCGG	GAGGAAGTCT
781	CGTGTCTTGG	CTCGTCTTC	TTCGAGGTCA	AGCCCAAGAC	CCTTGAGAAA	ATCGTCCGAT
841	TCCTCAAGCG	TCACAATGTC	GAATTGAGC	CCTACTTCGA	TGAAACAGCC	CTCGAGTCTA
901	TCGATGATAT	TATCACTTT	CTGGACGCGG	GCGCCCGAA	GGTGTGTTGTC	AAGACCGAGC
961	AGTTGGCCGA	CCTCTCCGCA	TATGGCTCCC	GCGTTGCC	CATTGTCACT	GGAAGCAGCG
1021	CTGCTTGT	TTCCCTCCGC	ACCGAGAGCG	GCCTTTGCT	CTCCGGCTTC	GATCAGACTG
1081	CCTCCGAGGC	TGCACAGTT	CTGGAGGAGG	CCAGAGACAA	GAAAATTAC	CCCTCTTC
1141	TCAAGCCC GT	TCCTGGGGCC	GATCTGAAC	AGTTCATCCA	GGTCGCCGCC	AAGGCTAACG
1201	CCATCCCCAT	CCTGCCATCC	ACTGGCTTGA	CAACAAAGAA	GGACGAGGCC	GGAAAGCTTG
1261	CCATCTCCAC	CATCCTCTCG	AGCGTCTGGA	AGTCTGACCG	TCCCAGATGGT	CTGCTCCCCA
1321	CGTGTGT	TGATGAGCAC	GACACTGCTC	TGGGTCTGGT	CTACAGCAGT	GCCGAGAGTG
1381	TGAACGAGGC	CCTCAGGACA	CAGACTGGT	TCTATCAGAG	CCGGAAGCGC	GGTCTCTGGT
1441	ACAAGGGTGC	TACTTCGGGA	GACACTCAGG	AGCTCGTCCG	CATCTCGCTT	GACTGCGATA
1501	ACGATGCTCT	CAAGTTGTC	GTGAAGCAGA	AGGGTCGTTT	CTGCCACCTC	GATCAGTCCG
1561	GCTGCTTGG	TCAGCTAAA	GGCCTTCCC	AGCTCGAGCA	GACTTTGATT	TCGAGGAAAC
1621	AGTCTGCC	CGAGGGCTCC	TACACTGCC	GTCTCTTCTC	CGATGAGAAG	CTAGTCGGG
1681	CCAAGATCAT	GGAGGAGGCT	GAGGAGCTCT	GCACCGCTCA	GACCCCCCAG	GAAATCGCCT
1741	TTGAGGCTGC	CGATCTTCTC	TACTTGCTC	TTACCAGGGC	CGTTGCTGCC	GGCGTTACTC
1801	TTGCGATAT	CGAAAGGAGC	CTTGACGCCA	AGAGCTGGAA	GGTCAAGCGC	AGGACTGGAG
1861	ATGCTAAGGG	TAAGTGGCT	GAGAAGGAGG	GCATCAAGCC	TGCGCGTCC	GCTCCCGCTG
1921	CCACTTCGGC	CCCTGTCA	AAGGAGGCCG	CCCAGGAGAC	CACCCCTGAG	AAGATCACCA
1981	TGAGACGTT	CGACGCC	AAGGTCTCTA	CCGAGGAGCT	CGATGCTGCT	CTCAAGCGTC
2041	CTGCGAAAA	GTGTCGGAT	GCATCTTACA	AGATCATTGT	CCCCATCATC	GAGGAGGTCC
2101	GCAAGAACGG	CGACAAGGCT	GTTCAGTCGT	ACACTCACAA	GTTCAGAGAAG	GCTACCTCTC
2161	TTACTAGCCC	CGTCCTGAAG	GCGCCCTTC	CCAAGGAGCT	TATGCAGCTC	CCTGAGGAGA
2221	CCATTGCTGC	CATCGACGT	TCCTTCGAGA	ACATCCGCAA	GTTCACGCC	GCCCAGAAGG
2281	AGGAGAACCC	CCTCCAGTC	GAGACCATGC	CCGGTGTGTT	CTGCAGCCGT	TTCTCTCGTC
2341	CCATCGAGGC	CGTGGCTGC	TACATCCCCG	CGGGTACCGC	CGTTCTCCCC	AGCACTGCC
2401	TTATGCTGGG	TGTTCCGCC	ATGGTCGCC	GCTGCAACAA	GATTGTTGTC	GCCTCTCC
2461	CCCGGCCGA	CGGAACCATC	ACTCCCAGA	TTGTCACGCT	CGCTCACAAAG	GTTGGGGCCG
2521	AGTCATCGT	GCTTGGCCG	GGTCCCAGG	CCGTAGCTGC	CATGGCC	GGCACCGAGA
2581	GCATCACCAA	GGTCGACAAG	ATTCCTGGCC	CCGGAACCA	GTTCGTC	GCTGCCAAAGA
2641	TGTTGTCAG	CAACGACACC	AACGCTGCC	TTGGGATTGA	CATGCCGCT	GGCCCGTCCG
2701	AGGTGCTGGT	CATCGCTGAC	AAGGACGCCA	ACCCCGCGT	CGTTGCTCG	GATCTCTGT
2761	CCCAGGCTGA	GCACGGCGTT	GACAGTCAGG	TCATCCTGAT	CGCTTAAAC	CTCGACGAGG
2821	AGCATCTTCA	GGCTATTGAG	GACGGAGGTT	ACCGTCAGGC	TATGGAGCTT	CCTCGCGTCC
2881	AGATTGTCCG	TGGCTCCATC	GCCCACCTCGA	TCACCGTGCA	GGTCAAGACC	GTCGAGGAGG
2941	CCATGGAGCT	CAGCAACAA	TACGCTCCTG	AGCACTTGAT	CCTCCAGATC	AAGGAGGCCG
3001	AGAAAAGCTG	CGATCTTGT	ATGAACGCTG	GTAGTGTCTT	CATTGGCGCT	TGGACTCCTG
3061	AGTCCTGTT	CGATTACT	GCTGGTGT	ACCACTCGCT	GCCTAAGTTA	CATATCATAA
3121	ATAGCCCCGC	TTCACAGATT	CTTCAGCTAA	CGTCAAGACA	CATAGCTACC	TATGGTTTTG
3181	GCAAGCAGTA	CTCTGGCGTC	AATCTCGCT	CGTTGCTCAA	GCACATTACC	AGCTCCAACT
3241	TGACTGCCGA	GGGTCTAAA	AACGTCGGCC	AGGCTGTCA	GCAGTTGGCT	AAGGTTGAGG
3301	AGCTCGAGGC	TCACAGAAGG	GCGGTCAGCA	TCCGTCTGA	GCACATGAGC	AAGAGCAACT

**Figure 7 continued**

3361 AGACGGAAAT TCTTTTCGA AGTTGAAAAA AAAACAAGAA CAAAAGGATG TAGTGGTTG  
 3421 ATGTATATCT GGGTCATTG GGGCACATAG AGTAATGATA ACGAGTTTG GACATTGTAC  
 3481 TGTTCTGTAC AGGCTGAAGA TCAGTACATG AATCTGTTGG TAAGTGTAGA GACCCAAACG  
 3541 TCCCCTGAGT TTTTCTCCCT GTTCCAGAGA GGTGCTCGTC CCTGGGTGTT TATTTTCATT  
 3601 ATTACATCAA CCTTTTATT TATTTTATT TTTATTTAC TTTTTTCC TTTTTTCAG  
 3661 ATCATCGCTA CATGAACGGG GGAAGCACAG ACGATCGAA CGTGGATGTC ACAATGTGCG  
 3721 TGCAGTGATG CTGCATTGCA TGAAGCGCCC ATCTCAATAT ACTTGCAGTC TTGCGCGTTG  
 3781 CACGTAACT TCCCAAACAA CGGAATAAAA GACGGCGAA AATGAAGATA AAAAAAAACC  
 3841 ATAATAAAA TCGGAGGGAG TGTGGGAAAT GGTTCTTT AGCATTAGA CCCCATAGCC  
 3901 GTGCACGCC GGGTACAGAC AGGTTCATCG ATGTTGACAT TGACTGGGAC ACCAGGTCTA  
 3961 TCTATTCAT CTCCTGTCT CTACCATACA TCGGGACATC GGACATCTCG CTGTACCCCC  
 4021 CACACCCACA AAGTCTTATA AAAGCGCCAC ACCCGAGGAG GTCGCGTCGG CCCCACGAAC  
 4081 TCCGTGCCTC CCTGCCTGTT TACAGGGACC GAACGCTGGA GAAGCTTAGT TTCTGACAT  
 4141 CGGGCCTTAC CGAGCAGGAA AAGGGACAGC TCATAGGCGA GGAGGGATT GAAGATGGGG  
 4201 ACATTTGGA TGATTGAGA GGAGGAACTA GGTACTGTAT CATGATAGTT CGGGGCAGCA  
 4261 TCTTGGCTGG GACATTGTTA ATACCTCGAT ATGATGAAGT GGGAGGGAGT TTTTCATGT  
 4321 CTTGCCAAG TCCCCTAAT CTTTTTTTT TTTTGTACCA ACACCCAAGA TTCGGAGAAT  
 4381 AGTGTAAAGGA TTGCAATTCA CAAGTGGAAAG TCTGAGGATC TTTTATATTC TTGTCCTTCC  
 4441 GCGGACTGTT AACGATCTA CAGCGAGCGA GCGAGCGGTC GGATGCGCTG ATCTGATAGG  
 4501 TGCAATATAC GGCGCTTTC TCCGGTGTAG TAGTGTAAAGC TCTGTCGGCA TAGTAGTACA  
 4561 CTAAAAAAC CCTTGCAATT CATGATCTGC TTGCTATTCA TTCCGAGTTA TTTCAGTGGT  
 4621 CACATTCGA GATTACACAGC CATCCATCCA TATGGAAAAA TCCATTCCCA TGCTTCCTCC  
 4681 CCCCCACTAT GTATGTGACC ACACGCTGCT GTCAGAAATGC CAACGGTCTC AGGTACCCCTC  
 4741 GTCCGACTGTT TTGGCATGGA GTTACATACA CTACTAGTGT AGCCCCGGGC CAAGCTACCC  
 4801 CGTCAAATCT ATACATATCT ATAATGGGTT TCAGGTGTT CGTTCGCTGT CAATCAAGTT  
 4861 TGAAACATCA CTGGGGCCGT TGGACGGTGT ATTAGACCAT TGGCTCCCTC AGCTGGCGC  
 4921 TGGCGGTTG GGTGGCAAT AACGGGACTG GACTTGAGAG GGACGAGGAG AGTCGGTTGG  
 4981 CTGCTTACAC TACACTACAA GCGTTCCAC CTAACCGACG AGTCCCGTT TCCATTGTG  
 5041 TGCCCTAACCT ATCATCTAGG GATGTCAGGG TTTGGCCGGA TCAGGGTATG TTTGGTTGAC  
 5101 TGTTGTCTG TCTGATTGGG TACATATCAT GGTAGGTGTC TCGAGAACAG TAGAGTACTC  
 5161 GGGCTTAGCG TTTGGATGAT TACCGCAGAT ATGAGTTGTA GGCCGCCATG CAGTTGCTTG  
 5221 CCCATAAGCA GAAAGTGTCT TGGGATATAT TTCTGCTCTT TCAAAGGTCA CGAGGTCCCTG  
 5281 GGACGAGCGG CATGCCATC CAAAGGGTTG AACATGAGAA ACCGGAATGG CCTTTGCGTT  
 5341 GAAATACAAA AAGTCAGAA TAAAATCGT TGAGGATAGG GACGGTGAAG CAAGCAAATA  
 5401 TGGTAAGGGA GGTACTGTA TGTAGGTGCT CAGCCTAACG CCAATTCTT GGCCCCAAAG  
 5461 CAGCAGTTG CTGTCAGTGC TGCTCGTGTG AGCCTGGTA GTGAAACCTA AACTGCTAAC  
 5521 ACAGCGCAAG TGCGCATGTA AAGATATTGT GGGAGGATCT GTATGGATGG ATGAGATTAC  
 5581 TGCTTGGTGT TGGTTGCGAG GCACTGCGGC TGTTAGGCTT TGCTGTGCC CGTTGACGA  
 5641 AGAAATACGC GGAACATATAA ATTGGATACC TAGACTTACT GCCTATGGGA GGTATCTACC  
 5701 GACGTAGCCG ACGGATTCTA GCAACATCCC GACTTGCTT GTAGTGTACT ATGATAGCAG  
 5761 CACAGTGGGG TGTGCTCTT TGTGAGCATG GGCTCTTTT TTTTTTTCC CCCTTCCTA  
 5821 GGGCGTTGAC TGGACTTGCT CTATCGTTCC CAAGGTAGGT GCGCGTCATG GATTTTCCCA  
 5881 AGCCGTCTCC CGCCAGATTG TCGTCATAGT GTCATGATGA CCTCGGTGCG TGGGGCTGCG  
 5941 TGGTTACGGG GAGCTGGGAC CGCTAGGCT CAGTGGTTGT GCCATTCTCAGG GTGGGTGTT  
 6001 GGAGTAGCGG TAGAGGCGCT TGGAAAGTTG GCTAGCGAA ACCCTGGAAT ATCTTGTACC  
 6061 CTTCGATTCC TTCTCGGGCT GCCCATGTGC TGAGGTGATG CCGGGGATCT GGCGCCAATC  
 6121 ATCCATTGAG GTTCCCGCAG CTTCCCGGTG CCGCGCCGG GCCCAGTTGC TCACAGGACA  
 6181 CACCTAGACG CAGGGGCACA GGGGCACCGT TTGGTGTGCA ACTGGGTACC TGGTAGCTGT  
 6241 AGCAAGCACT CCACCGTCTG TGCAATCCCC CAATCCACGG CAGGAACCTA GCACCGCCGC  
 6301 GGCACCGAGT GAGCGAATCC ATCCGATTG GATCCCAATT CTTGCCCTG CCATCCTTCT  
 6361 TTCTTCCCAC TTGGCGAAC CAAACATTCC TTGGCTCTG TGACTCGTGT TGATCTTCAC  
 6421 TCTCTTTTT TCTTGGGCGA CCGACTTTT ATATCCGTC TTGCTTCCCC CTGGCCGTTG  
 6481 TCGTCTTTC TACAACCTAC TTCCGTTCAT TATCCCCCTT CTTGGTGTGCG TCGAGGACCC  
 6541 AAAAACAGAA CAATTCCGGC TCTTCCAGGT GGCTTGGGT CGACTGTTA GCTCTGACC  
 6601 ACTAGCGCT TACCTCTCT TGATGTTAT ATTTGGATAT CATTGAACTA CTCTTCTTG  
 6661 AAACGGCAGA CGAACGGAAC AGTCCCTACG GTTTATTAGC GATATACGTT GTACTGATAT  
 6721 CCTGAGCAAG AAGAGGCAA TTATCAATT TGCACTCTCCTC ATCGTCGCTG CTCATCGCAG  
 6781 CTCCCTGCT CGCCAATGTA TCGGCCGAAC CGATTAGGAT ACCCCAAACGC GATGTTCTCC  
 6841 GTGGTATCAA CATCACAGCA ACTTGCCGTT CGAGCACTAC CGAATTGCGCC CAGCGGTGGA  
 6901 TATGCCCTG CCGTTGAGA CTGTCCTAACG ACCAAGCCGA CGCTCCGGAA GGCGGTGGAT  
 6961 TTGTCGAACG AGGAGAAGAA CTGGTTGTCG ATCCGGAGGA AGAACACCAT CCAGCCCATG  
 7021 AGGGACCTAC TGAAGAGGGC CAACATCACT GGGTTCGATT CGCAAACCTT CATGAATGAG  
 7081 GCCGCAACA ACGTCTCGCA ACTGCCCAAT GTCGCCATTG CCATTTCAGG AGGCGGCTAT  
 7141 CGTGCCTCA TGAACGGCGC CGGCTTCGTT GCTGCTGCGG ATAACCGGAT TCAAAATACC  
 7201 ACGGGCGCAG GTGGTATTGG AGGCTGTTG CAGTCCAGCA CATATTGTA TGAAAACCA  
 7261 TGCCTTCTTG TGGTTCTCT TATCTGTTT TCGAGTGTCA ACTGCGCCAG TTCGACGTTG

**Figure 7 continued**

7321 GCGGGCTGTG GACGACCTTG CTGGTGAACA TGTCTTGGAC TCCATGCCCT TTTTTCGTT  
7381 CCCTAAAATC CCAAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAATTGAG  
7441 GACCGTACT GTAAATTGCT AACGCAACTC TAGGGCCGGA CTTTCTGGTG GTGGCTGGCT  
7501 TGTCCGCAGT TTGTTCTCCA ACAACTTCAG TAGCATTGAG ACCCTGCTGA GCGAGAACAA  
7561 AGTCTGGAC TTTGAGAACT CCATCTTTAA AGGACCAAG GAGGCTGGCC TTAGTACTGT  
7621 CAACCGTATC CAGTACTGGT CGGAAGTGGC AAAGGAAGTT GCGAAGAAGA AGGATGCTGG  
7681 CTTCGAGACA AGTATAACAG ACTACTGGGG CCGAGCATTG AGTTACCAAC TGATCGGAGC  
7741 CGATATGGGC GGCCCAGGCTT ACACCTTCTC CAGCATTGCC CAGACCGACA ACTTCCAGAA  
7801 GGCGAAACG CCGTCCCTA TTCTGGTAGC TGACGGCCGC GCGCTGGAG ACACCACAT  
7861 CTCCCTCAAT GCTACCAACT ACGAGTTCAA CCCGTTCGAG ACGGGTAGCT GGGACCCGAC  
7921 CGTCTATGGC TTTGCGCCGA CCAAGTACCT CGGCGCCAAC TTCAGCAACG GCGTGATCCC  
7981 ATCGGGAGGC AAGTGCCTG AGGGTCTCGA CCAAGCCGGC TTCTGCATGG GCACCAGCAG  
8041 CACGCTCTC AACCAGTTCC TTTTGGCCAA CATCTCCAGC TAGCACGGTG TTGCCAGACG  
8101 TGCTCATCGA GGCCGTGACT TCTGCCTCA AGGAAATCGG CGCCAAGAGG ACGACGCTC  
8161 CCAAATCATC CTTAATCCGT TCCCTGGACTG GAACAACCGG ACCAACCCCA ACGCCGACAC  
8221 GCTCGAGCTC GACCTGGTCG ACGCGGGCGA AGATCTGCAG AATATTCCGC TCAACCCGCT  
8281 CACCCAACCC GTGCGCGCCG TCAGACGTAT TTCTCGTGTG CACTACGAG CGCACTTTCG GCTCTATTT  
8341 AAACCTGGCCC AATGGCACCG CCCTGCGCGC CACCTACGAG CGCACTTTCG GCTCTATTT  
8401 CAACGGGACA CTCTCCCT CGATCCCCGA CGACTGGACG TTTATAAACC TAGGCCTCAA  
8461 CAACCGCCCC TCTTCTTCG GCTGCGATGT TAAGAACCTT ACCTTGAACG CCAACCAAAA  
8521 GGTTCCCCCC TTAATCGTCT ATGTCACCAA CGCGCCCTAT ACCCGCGCTGA GCAACGTGTC  
8581 CACCTTCGAT CCGTCATACA CGATGTCTCA GCGCAACGAC ATCATCGCA ACGGATGGAA  
8641 CTCAGGCCAGC CAGGGAAACG GCACGCTGGA TTGGAGTGG CCCACTTGCG TCGCCTGCG  
8701 GGTTATCAGC AGGAGCTTAG ATCGGTTGGG CAGGCAGACG CCAGCCGGGT GCAAGACTTG  
8761 CTTTGACAGG TATTGCTGGA ATGGCACAGT GAACTCCAAA GATACGGGGG TTTACATGCC  
8821 TGAGTTCAAG ATTGCGGATG CGCATGCCCT GGACTCGGGT GCTGTTGCTA TCGGAAAGAT  
8881 GGTGAATGTC TGGTCGTCGG TTGTGGTGGG AGTTGTGGCG GCTACTTTGT TGTGTAAGGG  
8941 GTAGGGGAGA CGTGATGATA TTCCAGTCTG ATGAAGTTGA GACTGGACTG GAGATGCCA  
9001 AGGATGCGGA GGGAAAGGA TGCGGGTGT TAATGTCATG ATGGATGAAG AGTCATGGAT  
9061 CATGGAACGA CGGGGCGGGG ATATTGGATG ATGGATATAC CACACTGCAT GCATGCTCTA  
9121 TTGATAGTAT GCTTTGGCAT TTACGTTAA CAATCAATTG CTCCATCCTG ATGTTCTATC  
9181 TTTTTCGACA ATGGATTGAT ACTACTCCCTG TTGCTTCGCT CTTGAGGGTTG GAAGGACTTG  
9241 AGGITGGAAG GACTTGAGGT TGTTGTTCT GAGGGAGGGT ATCGAAGTAT CATCTGTGCT  
9301 GATGCCGATT GATAGACTGT CCTCTTCTTC GAGGCAACGA ACGGTCGGAT GAGCCTCTT  
9361 AATCATGATG CTCAGTGCCA CAAAAAGGCT CCAGCACAGC TGCCCCACACC TTTCTTGCT  
9421 CGCCGTTCTC TCCCTTTCT TTTCCTCTGT TTCCCTTCTT CCTTCCATC TCATCCCGTA  
9481 CCAGAGTGCC CACCGGGTAT ATATATTACC TCCCTGGCCG TTCTCCTTTG ACCAATAAAT  
9541 CGCTTGGTCG AGTGGCGTAA CGGTTTACCG TCTACACTTA TCACTCAAAC CAAACCAAAC  
9601 CATCGAAGAA GTGACCTATC GTTTCGAGGG AACGGTGATG TTCTTACGAC CAAGTTAAC  
9661 CAAAGAGCGT TCCACATCGT TGAACCGTCT CCTCCAGTTG GATCTGTTA ACTTCCGCAG  
9721 CGACTGAAGA AGGTATCACT TTTTTTTGG TTCCAAAAAA AAAAAAAA ATTAC

**Figure 8** Nucleotide sequence of the *his-3 cog<sup>E</sup> lpl* region of linkage group I in the StLawrence wild type strain of *Neurospora crassa*. This differs from that in the Lindegren strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the weak recombinator *cog<sup>E</sup>* and also the remnant of a transposable element *Guest* within the replaceable sequence 3' of *his-3*. StLawrence strains carry *rec-2<sup>+</sup>* which prevents the initiation of recombination at *cog*.

1	ACCGGGAATC	GTAGCGGGCG	CTAAGGCCAA	GCCGCGGCAC	GGGTCACTGA	CCCAATGCAG
61	CGCATTCCGT	CAGCAACTGA	AGTGGATGTA	CAAGTACATA	GTAGTAGATC	GCAACTGGAG
121	ATCACTCGCA	CCGTGCCGCA	GAACAAGGGC	GACGAGCCTC	AGGGCAGTTT	AGCCTGCCGT
181	AACAGCACAG	ACCATAGCTT	ATTTCACCT	GGGCGGGCGG	GCGACGGCGG	CACTGACATC
241	GGCAAGGCAG	CATCAAGCAA	CCCCTCTGTT	GCTTGCCAGC	TGCAGGCCAA	CGTCAGCGGT
301	ACAAGGAGAA	ATCTGGAAGG	AAAGACTCT	GGCACCGACA	GGATGGCACG	CGGGAAAAGT
361	TCCCAATGCA	TGAGATGAGG	GGCATTTGCA	TTGCTCTCCG	TCACCCAGTG	CGAACCCCAA
421	CCCCAACCATCA	GCGTCTGTCG	ATACATGGAG	CGCGAAGTCG	AGAAACCTGT	AATTCTGGT
481	AACTTTCAGG	TACACAGTAC	GTACTGATCC	TGGTATCAA	CCTTGCCTGC	CGAGTTTCG
541	ACGGAAAGAG	GTGTGAATTG	TGAAAGAGAT	ATACCAAAT	ACCCGATTTT	CATAAAGCCC
601	GAGTCTTTC	TGTACATAAG	CGACACTCGA	AGCGGGCCTC	ATCTTCATAG	CCTGATAGCT
661	TGTAATACTC	CATCCTCGTA	TCTCACTTGA	CCTTGAGTT	AACCCACGT	CAAACCTCAC
721	CCGACACATC	GACGGATTGG	GGAACACGCAC	AATACTGAA	AAGCGAGAAA	ACCAAACAGA
781	GGAAAACACC	ATGGAGACAA	CACTTCCCT	CCCCCTCCTC	GTGGTGTGCA	GTGTTCTCC
841	CGGACTGAAT	SACATCAAGG	AGGGCCTCAG	CGGGGAGGAA	GTCTCGTGTC	TTGGCTCGT
901	CTTCTTCGAG	GTCAAGCCCC	AGACCCCTGA	GAAAATCTG	CGATTCTCTA	AGCGTCACAA
961	TGTCGAATT	SAGGCCCTACT	TCGATGTAAC	AGGCCCTCGAG	TCTATCGATG	ATATTATCAC
1021	TCTTCTGGAC	GCGGGCGCCC	GCAAGGTGTT	TGTCAAGACC	GAGCAGTTGG	CCGACCTCTC
1081	CGCATATGGC	TCCCGCGTTG	CCCCCATTGT	CACTGGAAGC	AGGGCTGCTT	TGCTTCTCTC
1141	CGCCACCGAG	AGCGGCCTT	TGCTCTCCGG	CTTCGATCAG	ACTGCCTCCG	AGGCTGCACA
1201	GTTTCTGGAG	GAGGCCAGAG	ACAAGAAAAT	TACCCCTTC	TTCATCAAGC	CCGTTCTGG
1261	GGCCGATCTC	GAACAGTTCA	TCCAGGTCGC	CGCCAAGGCT	AACGCCATCC	CCATCCTGCC
1321	ATCCACTGGC	TTGACAACAA	AGAAGGACGA	GGCCGGCAAG	CTTGCATCT	CCACCATCCT
1381	CTCGAGCGTC	TGGAAGTCTG	ACCGTCCCAG	TGGTCTTCTC	CCCACCGTTG	TCGTTGATGA
1441	GCACGACACT	GCTCTGGTC	TGGTCTACAG	CAGTGCAG	AGTGTGAACG	AGGCCCTCAG
1501	GACACAGACT	GGTGTCTATC	AGAGCCGGA	GCGCGGTCTC	TGGTACAAGG	GTGCTACTTC
1561	CGGAGACACT	CAGGAGCTCG	TCCCACATCTC	GCTTGACTGC	GATAACGATG	CTCTCAAGTT
1621	TGTCTGAAAG	CAGAAGGGTC	GTTTCTGCA	CCTCGATCAG	TCCGGCTGCT	TTGGTCAGCT
1681	CAAAGGCCTT	CCCAAGCTCG	AGCAGACTTT	GATTTCGAGG	AAACAGTCTG	CCCCCGAGGG
1741	CTCCTACACT	GCCCCTCTCT	TCTCCGATGA	GAAGCTAGTC	CGGGCCAAGA	TCATGGAGGA
1801	GGCTGAGGAG	CTCTGCACCG	CTCAGACCCC	CCAGGAAATC	GCCTTGAGG	CTGCCGATCT
1861	CTTCTACTTT	GCTCTTACCA	GGGCCGTTGC	TGCCGGCGTT	ACTTTGCCG	ATATCGAAAG
1921	GAGCCTTGAC	GCCAAGAGCT	GGAAAGTCAA	GCGCAGGACT	GGAGATGCTA	AGGGTAAGTG
1981	GGCTGAGAAG	GAGGGCATCA	AGCCTGCGGC	GTCCGCTCTC	GCTGCCACTT	CGGCCCCGT
2041	CACCAAGGAG	GCCGCCAGG	AGACCACCCC	TGAGAAGATC	ACCATGAGAC	GTITCGACGC
2101	CTCCAAGGTC	TCTACCGAGG	AGCTCGATGC	TGCTCTCAAG	CGTCTCGCG	AAAAGTCGTC
2161	CGATGCCATC	TACAAGATCA	TTGCCCCAT	CATCGAGGAC	GTCCGCAAGA	ACGGCGACAA
2221	GGCTGTTCTG	TCGTACACTC	ACAAGTTCGA	GAAGGCTACC	TCTCTTACTA	GCCCCGTCT
2281	GAAGGCAGCC	TTCCCCAAGG	AGCTTATGCA	GCTCCCTGAG	GAGACCATTG	CTGCCATCGA
2341	CGTGTCTTC	GAGAACATCC	GCAAGTTCCA	CGCCGCCAG	AAGGAGGAGA	AGCCCCCTCCA
2401	GGTCGAGACC	ATGCCCGGTG	TTGCTGCGAG	CGTTTCTCT	CGTCCCATCG	AGGCCGTCGG
2461	CTGCTACATC	CCCGGCCGTA	CCGCCGTTCT	CCCCAGCACT	GCCCTTATGC	TGGGTGTTCC
2521	CGCCATGGTC	GCCGGCTGCA	ACAAGATTGT	GTTCGCTCTC	CCTCCCCGCG	CCGACGGAAC
2581	CATCACTCCC	GAGATTGTC	ACGTCGCTCA	CAAGGTTGGG	GCCGAGTCCA	TCGTGCTTGC
2641	CGGCCGTGCC	CAGGCCGTAG	CTGCCATGGC	CTACGGCACC	GAGAGCATCA	CCAAGGTCGA
2701	CAAGATTCTC	GGCCCCCGTA	ACCAGTTCTG	CACTGCTGCC	AAGATGTTCG	TCAGCAACGA
2761	CACCAACGCT	GCCGTTGTA	TTGACATGCC	CGCTGGCCCG	TCCGAGGTGC	TGGTCATCGC
2821	TGACAAGGAC	GCCAACCCCG	CGTTCTGTC	CTCGGATCTC	CTGCTCCAGG	CTGAGCACGG
2881	CGTTGACAGT	CAGGTACATCC	TGATGCTAT	TGACCTCGAC	GAGGAGCATC	TTCAGGCTAT
2941	TGAGGACGAG	GTTCACCGTC	AGGCTACGGA	GCTTCCTCGC	GTCCAGATTG	TCCGTGGCTC
3001	CATCGCCCAC	TCGATCACCG	TGCAGGTCAA	GACCGTCGAG	GAGGCCATGG	AGCTCAGCAA
3061	CAAGTACGCT	CCTGAGCACT	TGATCCTCCA	GATCAAGGAG	GCCGAGAAGG	CTGTCGATCT
3121	TGTCTGAAAC	GCCGGTAGTG	TCTTCATTGG	CGCCTGGACT	CCTGAGTCCG	TTGGCGATTA
3181	CTCTGCTGGT	GTAAACCACT	CGCTGCGTAA	GTTACATATC	ATAAAATAGCC	CCGCTTCACA
3241	GATTCTTCTG	CTAACGTCAA	GACACATAGC	TACCTATGGC	TTGGCAAGC	AGTACTCTGG

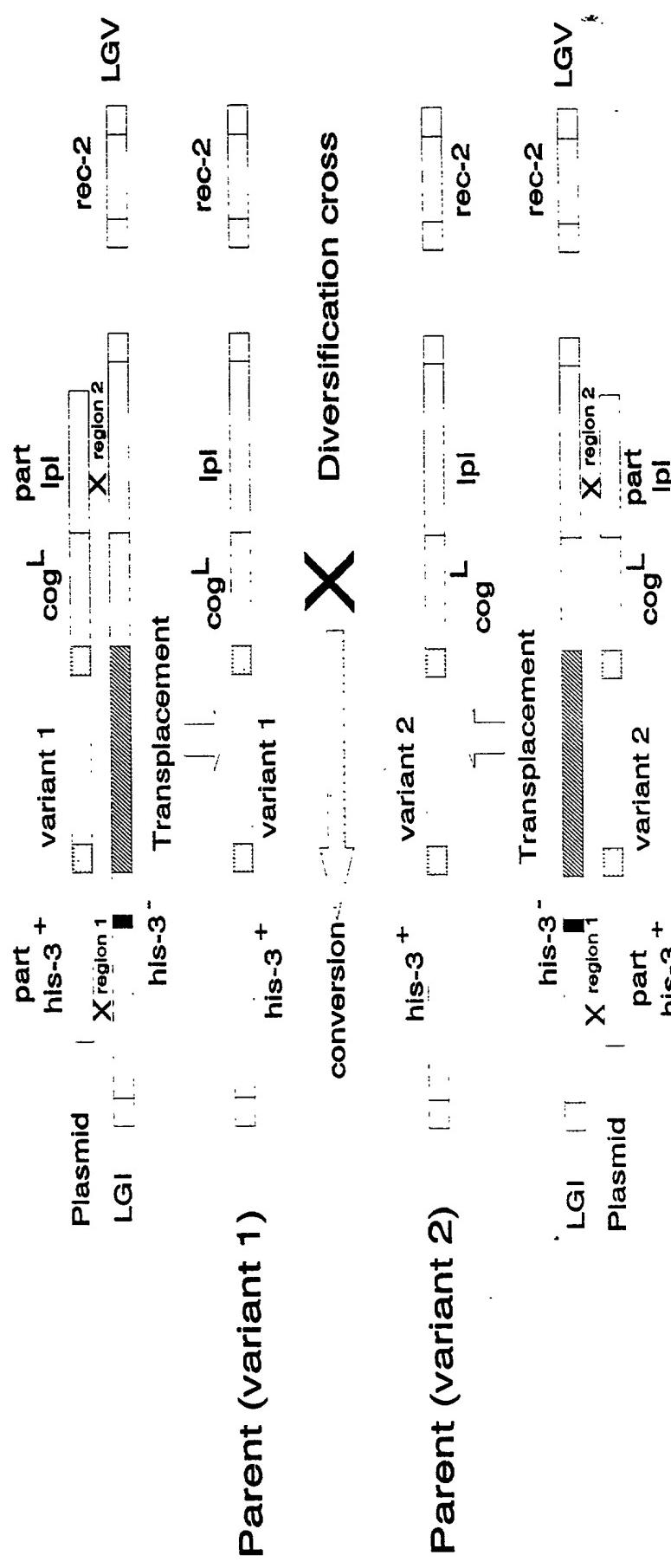
**Figure 8 continued**

3301	CGTCAATTTC	GCCTCGITCG	TCAAGCACAT	TACCAGCTCC	AACTTGACTG	CCGAGGGTCT
3361	CAAAAACGTC	GGCCAGGCTG	TCATGCAGTT	GGCTAAGGTT	GAGGAGCTCG	AGGCTCACAG
3421	AAGGGCGGTC	AGCATCCGTC	TTGAGCACAT	GAGCAAGAGC	AACTAAACGG	AAATTCTTT
3481	CGAAGTAGCA	AAAAAAAAAA	AAAAAAACAA	GAACAAAAGG	ATGTAGTGGG	TTGATGTATA
3541	TCTGGTCAT	TTTGGGCACA	TAGAGTAATG	ATAACGAGTT	TTGGACATTG	TACTGTTCTG
3601	TACAGGCTGA	AGATCAGTAC	ATGAATCTGT	TGGTAAGTGT	GGAGACCCAA	ACGTCCCTTG
3661	AGTTTTTCTC	CCTATTCCAG	AGGTGCTCGT	CCCTGGGTGT	TTATTTCAT	TATTACATCA
3721	ACCTTTTTT	TTTTTTTTT	TTTTTCAGAT	CATCGTACA	TGAACGGGGG	AAGCACAGAC
3781	GATCGAAACG	TTGGATGTCAC	AATGTGCTG	CAGTGTGATC	GCATTGCATG	AAGGCCCAT
3841	CTCAATATAC	TTGCAGTCTT	GCACGTTGCA	TGTGAACTT	CCAAACAACC	GAATAAAAAGA
3901	CGCGAAAAAA	TGAAGATAAA	AAAAAAACAT	AAAAAAATC	AGAGGGAGTG	TGGGAAATGG
3961	TGTCTTTAG	CATTCAAGACC	CCATAGCCGT	GCACGCCCG	GTACAGACAG	GTTCATCGAT
4021	GTTGACATTG	ACTGGGACAC	CAGGTCTATC	TATTTTATCT	CCTGTCCTCT	ACCATACATC
4081	GGGACATCGG	ACATCTTGCT	GTACCCCCCA	CACCCACAA	GCCTTATAAA	AGCGCCACAC
4141	CCGAGGAGGT	TCGGTCGGCC	CCACGAACTC	TGTGCTCTCC	TGCCGTGTTA	CAGGGACCGA
4201	ACGCTGGAGA	ATCTTACTAG	TTTCTGACCA	TCCGGCCTAC	CCGAGCAGGA	AAAGGGACAG
4261	CTCATAGCG	AGGAGGGATT	TGAAGATGGG	AACATTTGG	GTGATTGAG	AGGAGGAAC
4321	AGGTACTGCA	TCATGTAGT	TCGGGGCAGC	ATCTTGGTG	GGACATTGTT	AATACCTCGA
4381	TATGATGAAG	TAGGAGGGAG	TTTTTGCCTG	TCTTGGCAA	GTCCAGAGAT	CTGTTTATT
4441	TTATTTTTA	TGGATGTAGT	GTATCAACAC	CCAAGATTG	GAGAATAGTA	CTAGGATTG
4501	CATTACAAG	TGGAAGTCTT	GAGAATCGTT	GTATATCCTT	GTCTTCCTCG	GAATGTTAAC
4561	AATCTACAG	CGAGCGAGCG	AGCGGTCGGA	TGCGCTGATC	TGATAGGCGC	AATATACGGC
4621	CGCTTCTCC	GGTCGTGTA	TGTAAGCTCT	GTGGGCATAG	TACACTAAAA	AAACCCATTG
4681	ATTTCATGAT	CTGCCTGCTA	TTCAATTCCGA	GCTATTTCAG	TGGTCACATT	TCGAGGAAGA
4741	AAGAAAGCAA	CTAAGATTCA	CAGCCATCCA	TCCATCCATA	TGGAAGAATA	ATCCATTCCC
4801	ATGTTCCCTC	CCCCCCACTA	TGTATGTGAC	CACACGCTGC	TGTCAAGATG	CCAACGGCT
4861	CAGGTACCC	CGTCGCGACTG	TTTGGCATGG	AGTTACATAC	ACTACTAGTG	TAGCCCCGGG
4921	CCAAGTACCC	CCSTCAAAAT	TATACATATC	TATAACGGGT	TTCAAGGGTT	TCGTTCGCTG
4981	TCAATCAAGT	TTGAAACATC	ACTGGGGCCG	TTGGACGGTG	TATTAGACCA	TTGGCTCCCT
5041	CAGCTGTTG	GGGGCTGGC	GGCTGGGTCA	AACGGCAATA	ACGGGACTCG	AGAGGGACGA
5101	GGAGAGTCGG	TTGGCTGGCT	GCAATACAAG	CGTTCCCACC	TAACCAACGA	GTCCCGBTTT
5161	CCATTGTGT	GCCTAACCAT	CATCTAGGGA	TGTCAGGGTT	TGGCCGGATC	AGGGTATGTT
5221	TGGTTGACTG	TTGTCATGTC	TGATTGGTA	CATATTATGG	TAGGTGTC	GAGAACAGTA
5281	GAGTACTCGG	GCCTAGCGT	TGGATGATTA	CGCGAGATAT	GAGTTGTTGGG	CCGCCATGCA
5341	GTTGCTTGTC	CATAAGCAGA	AGTTGCTTTG	GGATATATTT	CTCGCTTTTC	AAAGGTACCG
5401	AGGTCTGGG	ACGAACGGCA	TCGCCATCCA	AAGGGTTGAA	CATGAGAAAC	CTGAATGGCC
5461	TTTGGTTGTA	AATAACAAAAA	GTCAAGAACAA	AAATCCTTG	AGGATAGGGAA	CCTGGAAAGCA
5521	AGCAAATATG	TTAAGAGAGG	TATACATCAA	CCCTGGTCA	ATTGTTAGCG	TGGTTCTTCC
5581	TCCACGTCT	CGTTCATGAC	GGTAAACAGT	ACCAGGCTAA	CAATTAAACC	AGGGTTGATG
5641	TGTACTGATA	TGTAGGTGCT	CAGCAAAC TG	CCAATTCTT	TGGCCCCAAG	CAGCAGTTTG
5701	CTGTCAGTGC	TGCTCGTGT	AGCCTTGGTA	GTGGAACCTA	AACTGCTAAC	ACAGCGAAC
5761	TGCGCATGTA	AAGATATTGT	GGGAGGATCT	GTATGGATGG	ATGAGATTAC	TGCTTGGTGT
5821	TGGTTGCGAG	GCAGTGGCC	TGTTAGGCTT	TGCTGTGCC	CGTCGACGA	AGAAATACGC
5881	GGAACATAAA	ATTGGATACCC	TAGACTTACT	GCCTATGGGA	GGTATCTACC	GACGTAGCCG
5941	ACGGATTCTA	GCAACATTC	GACTTTGCTT	GTAGTGTACT	ATGATAGCAG	CACAGTGTG
6001	CTCCTTGTGA	GAATGGGCTC	TTTTTTTTT	TCCCCCTTCC	CTAGGGCGTT	GACTGGACTT
6061	GCTCTATTGT	TCCCAAGGT	GGTGCCTCGT	ATCGATTTC	CCAAGTCTCC	CGCCAGATTG
6121	TCGTCATAGT	GTCTGATGA	CCTCGGTCGC	TGGGGCTGCG	TGGTTACGGG	GAGCTGGGAC
6181	CGCTAGGCCT	CAGTGGTTGT	GCCATTTCAGC	GTGGGTGTGT	GGAGTAGCGG	TAGAGGCCT
6241	TGGAAGTTGT	GCTAGCGGAA	ACCCCTGGAAT	ATCTTCTACC	CTCGATTCCCT	TCTCGGGCTG
6301	CCCATGTGCT	GAGGTGATGC	CGGGGATCTG	GCGCAATCA	TCCATTGAGG	TTCCCGCAGC
6361	TTCCCGGTG	CGCGCGCGGG	CGCAGTTGCT	CACAGGACAC	ACCTAGACGC	AGGGGACAG
6421	GGGACCGGTT	TGGTGTGCAA	CTGGGTACCT	AGCTGTAGCA	AGCACTCCAC	CGTCTGTGCA
6481	ATCCCCAAAT	CCACGGCAGG	AACCTCGCAC	CGCCGCGGCA	CCGAGTGGAGC	GAATCCATCC
6541	GCATTGGATC	CCAAATTCTG	CCCTTGCCT	CCTTCTTTCT	TCCCCACTTGG	CGCAACCAAC
6601	ACTTCCCTTG	GTCTGGGTAC	TCGTGTTGAT	CTTCACTCTC	TTTTTTCTT	GGGCGACCGA
6661	CTTTTTATAT	CCGTCCTTGC	TTCCCCCTGG	CCGTTGTCGT	TCTTCTACA	ACTACCTTCC
6721	GTTCATTATC	CCCTTTCTG	TTTCGGTCTGA	GGACCCAAAA	ACAGAACAAAT	TCCGGCTCTT
6781	CCAGGTGGCT	TGGGTGCGAC	TGTTTAGCTC	TTGACCACTA	GCCGCTTAC	TTCTCTTGAT
6841	GTTTTTATTT	GGATATCATT	AAACTACTCT	TTCTTGAAC	GGCAGACGAA	CGGAACAGTT
6901	CCTACGGTAT	ATTAGCGATA	TACGTTGTCAC	TGATATTCTG	AGCAAGAAGA	GGCAAATTAT
6961	CAATTATGCA	TCTCCCTTCG	TCGCTGCTCA	TCGCAGCTCC	CTTGCTCGCC	AATGTATCGG
7021	CCGAACCCAT	TAGGATACCC	CAACCGCATG	TTCTCCCGTG	TATCAACATC	ACAGCAACTT
7081	GCCGTTCGAG	CACTACCGGA	TTCCGGCCAGC	GGTGTGATATG	CCCCCTGCCGT	TGTAGACTGT
7141	CCCAAGACCA	AGCCGACGCT	CCGGAAGGCC	GTGGATTGTT	CGAACGAGGA	GAAGAACACTGG
7201	TTGTCGATCC	GGAGGAAGAA	CACCATCCAG	CCCATGAGGG	ACCTCCTGAA	GAGGGCCAAC

**Figure 8 continued**

7261 ATCACTGGGT TCGATTCCGA GACATTTATG AATGAGGCCG CCAACAACAT CTCGCAACTG  
7321 CCCAATGTCG CCATTGCCAT TTCAAGGAGGC GGCTATCGTG CCCTCATGAA CGCGGCCGGC  
7381 TTCGTTGCTG CTGCGGATAA CCGAATTCAA AATACCACGG GCGCAGGTGG TATTGGAGGC  
7441 TTGTTTGCAGT CCAGCACATA TTTGTATGTA AAATGGTTCT TCTTATCTCG TTTTCGAGTG  
7501 TCAACTGCGC CAGTTCAGAG TTGGGCGGCT GTGGACGACC TTGCTGGTGA ACATGCTTG  
7561 GACTCCATGC CCCTTCTTCG TTTCTCTAAA TCAAGAACGTC GAGGACCGTG ACCGTAATC  
7621 GCTAACGCAA CTCTAGGGCC GGACTTTCTG GTGGTGGCTG GCTTGTGCGC AGTTGTTCT  
7681 CCAACAACTT CAGCAGCATT GAGACCCCTGC TGAGCGAGAA CAAAGTCTGG GACTTTGAGA  
7741 ACTCCATCTT TAAAGGGCCC AAGGAGGCTG GCCTTAGTAC TGTCAACCGC ATTCACTACT  
7801 GGTCCGAAGT GGCAAAGGAA GTTGCCAAGA AGAAGGATGC TGGCTTCGAG ACAAGTATAA  
7861 CAGACTACTG GGGCCGAGCA TTGAGTTACC AACTGATCGG AGCCGATATG GGCGGCCGG  
7921 CTTACACCTT CTCCAGCATT GCCCAGACCG ACAACTTCA GAAGGCGGAA ACGCCGTTCC  
7981 CTATTCTGGT AGCTGACGGC CGCGCGCTG GAGACCCAT CATCTCCCTC AATGCTACCA  
8041 ACTACGAGTT CAACCCGTTG GAGACGGGTG GCTGGGACCC GACCGTCTAT GGCTTGCGC  
8101 CGACCAAGTA CCTCGGCGCC AACTCAGCA ACAGCGTGT CCCATGGGA GGCAAGTGC  
8161 TTGAGGGTCT CGACCAAGCC GGCTTCGTCA TGGGCACAG CAGCACGCTC TTCAACCAAGT  
8221 TCCTTTGGC CAACATCTCC AGCTACGACG GTGTTGCGG ACGTGCTCAT CGAACGCGT  
8281 ACTTCTGTCC TCAAGGAAAT CGGCGCCAAG AGGACGACGT CTCCCAAATC ATCCCTAATC  
8341 CGTCCCTGGA CTGGAACAAAC CGGACCAACC CCAACGCGA CACGCTCGAG CTCGACCTGG  
8401 TCGACGGCGG CGAAGATCTG CAGAATATTG CGCTCAACCC GCTCACCCAA CCCGTGCGC  
8461 CCGTGGACGT CATCTTCGCT GTGCACTCGT CCGCCGACGT GACAAACTGG CCCAATGGCA  
8521 CCGCCCTGCG AGCCACCTAC GAGGGCACTT TCGGCTCTAT TTCCAACGGG AACTCTTCC  
8581 CCTCGATCCC CGACGACTGG ACGTTTATAA ACCTAGGCT CAACAACCGC CCCTCTTTCT  
8641 TCGGCTGCGA TGTTAAGAAC TTACCTTGA ACGCCAACCA AAAGGTTCCC CCCTTAATCG  
8701 TCTATGCTCC CAACCGGCC TATACCGCGC TGAGCAACGT GTCCACCTTC GATCCGTCAT  
8761 ACACCATGTC TCAGCGAAC GACATCATCG GCAACGGATG GAACTCAGCC ACGCAGGGAA  
8821 ACGGCACGCT GGATTGGAG TGGCCCCATT CGCTCGCTG CGGGTTATC AGCAGGAGCT  
8881 TAGATCGTT GGGCAGGCAG ACGCCAGCC CGTGAAGAC TTGCTTGAG AGGTATTGCT  
8941 GGAATGGCAC AGTGAACCTA AAAGATAACAG GGGTTTACAT GCCTGAGTTC AAGATTGCGG  
9001 ATGCCCATGC CCTGGACTCG GGTGCTGTTG CTATCGAAA GATGGTGAAT GTCTGGCGT  
9061 CGGTTGTGGT GGGAGTTGTG GCGGCTACTT TGTTGTTGTA GGGTAGGGG AGACGTGATG  
9121 ATATTCCAGT CTGATGAAGT TGAGACTGG A TGGAGATCG CCAAGGATGC GGAGGGAAAG  
9181 GAATGCGTGG TGTTAATGTC ATGATGGATG AAGGGTCATG GATCATGGAA CGACGGGCG  
9241 GGGATAATTGG ATGATGGATA TACCAACTG CATGCATGCT CTATTGATAA TATGCTTTGG  
9301 CATTACGTT TAACAATCAA TTGCTCCATC CTGATGTTCT ATCTTTCGAC ACTGGATTGA  
9361 TACTACTCCT GTTGCTTCCC TCTTGAAGTT GGAAGGACTT GAGGTTGGAA GGACTTGAGG  
9421 TTGTTTGTTC TGAGGGAGGT TATCGAAGTA TCATCTGTG TGATGCCGAT CGATAGACTG  
9481 CCCTCTTCTT CGAGGCAACG AACGGTCGGA TGAGCCTTA ATCATGATGC TCAGTGCCAC  
9541 AAAAAGGCTC CAGCACAGCT GCCCCACACCT TTTTGCCCTC GTGCTCCCTT CTTTTTTTTC  
9601 CCCCCCTTTC TTCCCTTCCA TCTCATCCCG TACCAAGGTG CCCACCGGGT ATATATATTA  
9661 CCTCCCTGGC CGTTCTCCCT TGACCAATAA ATCGCTTGGT CGAGTGGCGT AACCGTTTAC  
9721 CGTCTACACT TATCACTCAA ACCAAACCAA ACCATCGAAG AAGTTACCTA TCGGTTCGAG  
9781 GGAACGGTGA TGTTCTTACG TTCAAGTTAA CCCAAAGAGC GTTCCACATC GTTGAACCGT  
9841 CTCCCTCCAGT TCTTGGATCT GTTAACTTC CGCAGCGACT GAAGAAGTAA TCACTTTTTT  
9901 TTTTTTGTG TCCAAAAAAA AAAAAAAA TTAC

**Figure 9** Construction of the components of the sequence diversification cross: Parent (variant 1) and Parent (variant 2). For convenience, plasmid sequences are shown as linear. The cross hatched region in the chromosome is dispensable. Stippled sequences in the plasmid indicate the multiple cloning site for inserting foreign DNA. Crossovers in region 1 and region 2 insert the foreign sequence to be diversified into chromosome 1 of *Neurospora crassa* adjacent to the recombination hotspot *cog*. Parent (variant 2) containing a version of the foreign sequence with multiple differences from that in parent (variant 1) is similarly constructed. Parent (variant 1) and parent (variant 2) are crossed and conversion events (stippled arrow) initiated (X) at *cog<sup>L</sup>* recombine the sequence differences in variant 1 and variant 2 to form new combinations. Sequences are identical (X) at *cog<sup>L</sup>* to distinguish variant 1 and variant 2. *rec-2* on linkage group V permits *cog<sup>L</sup>* to be active. For simplicity, genes not directly related to the diversification are omitted. See text for further details.



**Figure 10** Construction of parent (variant 1) and parent (variant 2) enabling selection of progeny that have experienced conversion in the foreign DNA. Complementing pairs of *his-3* alleles are used to obtain parent (variant 1) and a different pair of complementing *his-3* alleles are used to obtain parent (variant 2) as explained in the text. Parent (variant 1) and parent (variant 2) are crossed and his<sup>+</sup> recombinants are selected. These must all have experienced conversion events affecting the foreign DNA since the events begin at *cog*<sup>L</sup>. The *his-3* alleles in parent (variant 1) and parent (variant 2) are non complementing to ensure that selection yields recombinants and not aneuploid progeny having two copies of all or part of linkage group 1.

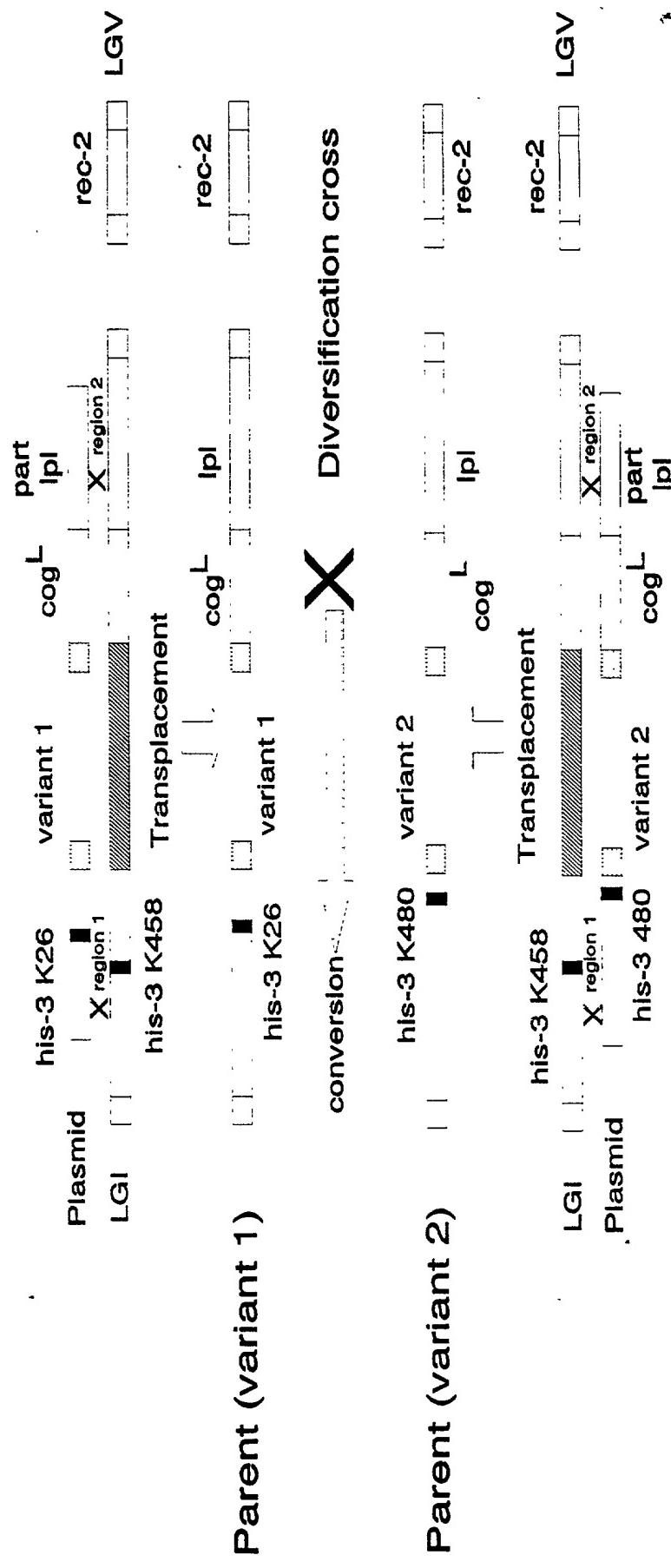
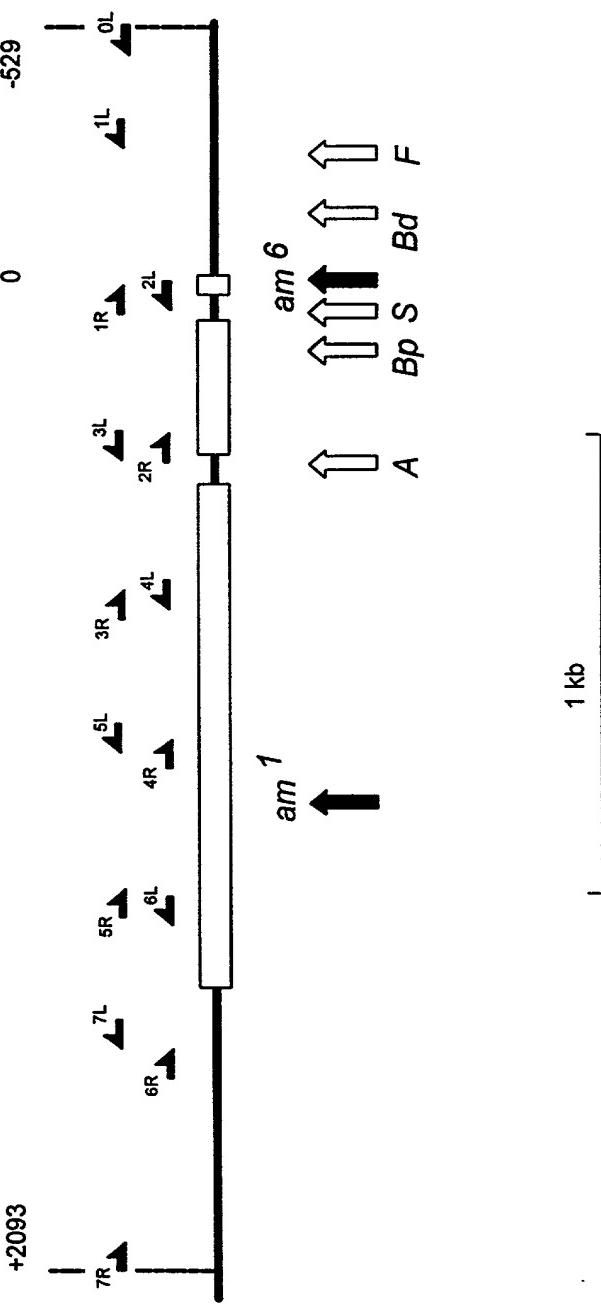


Fig 11



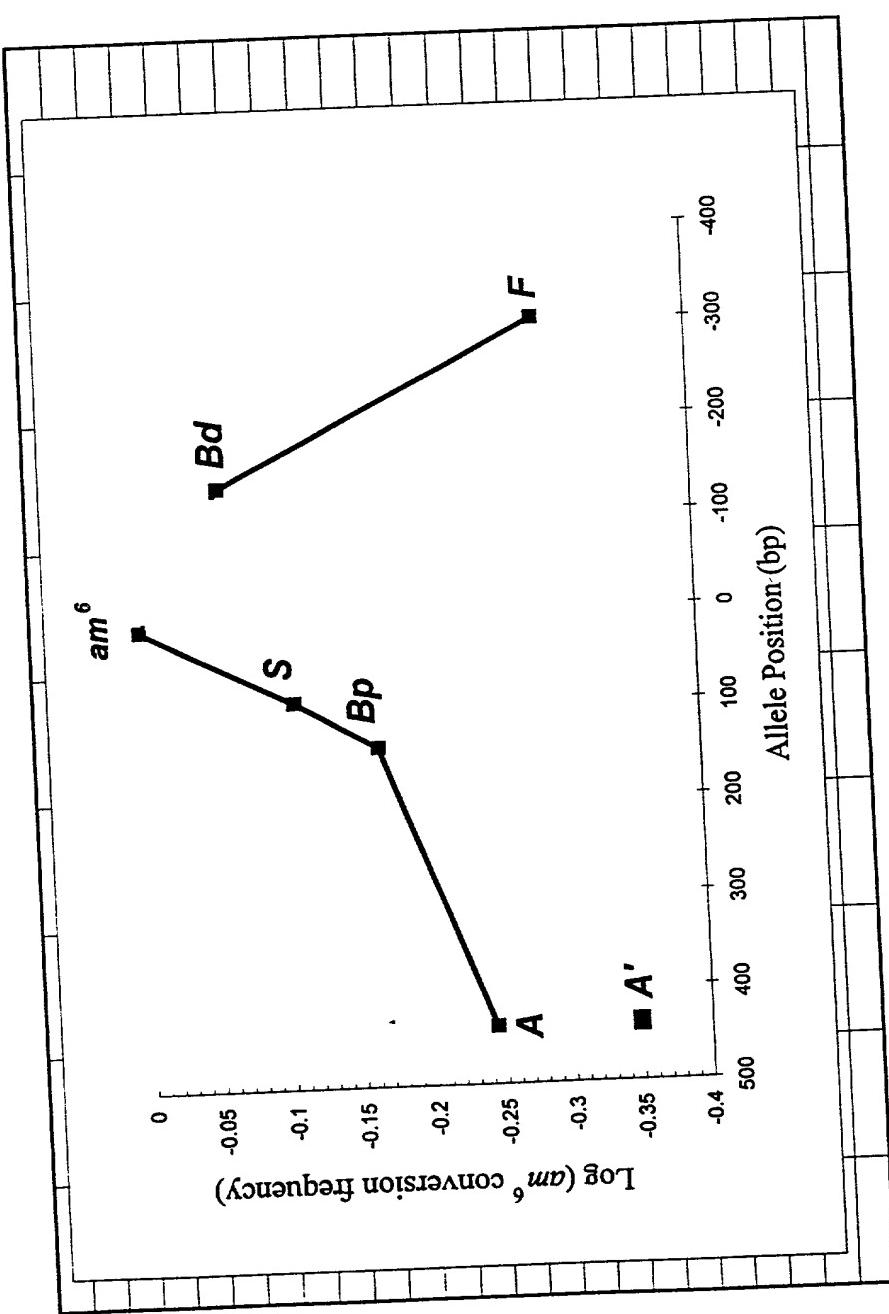


Fig 13

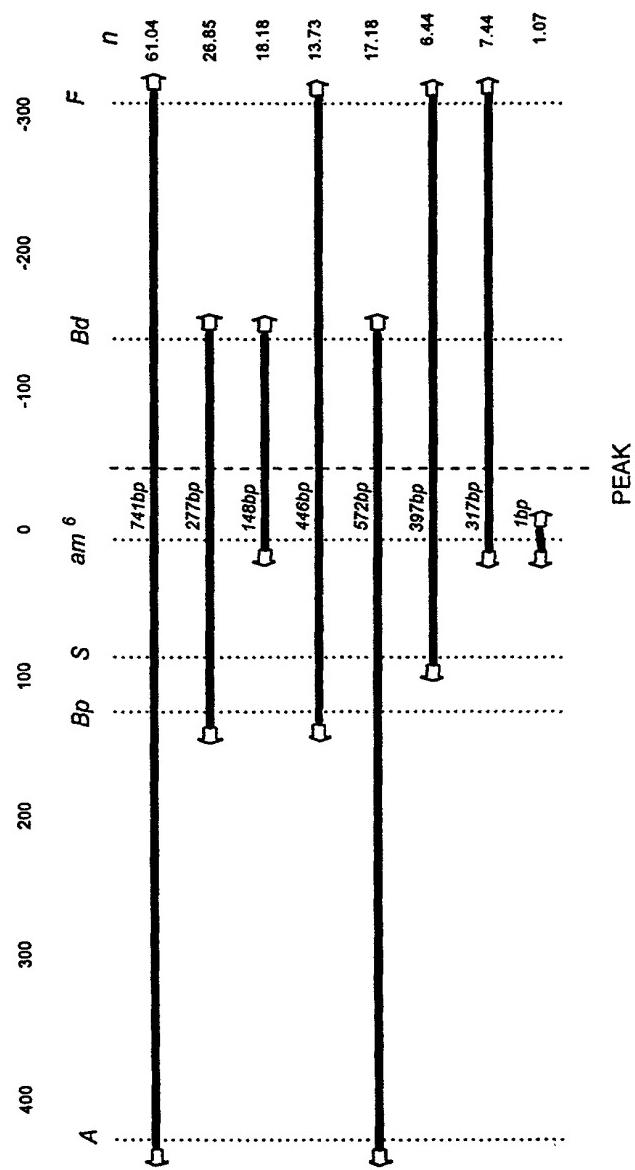
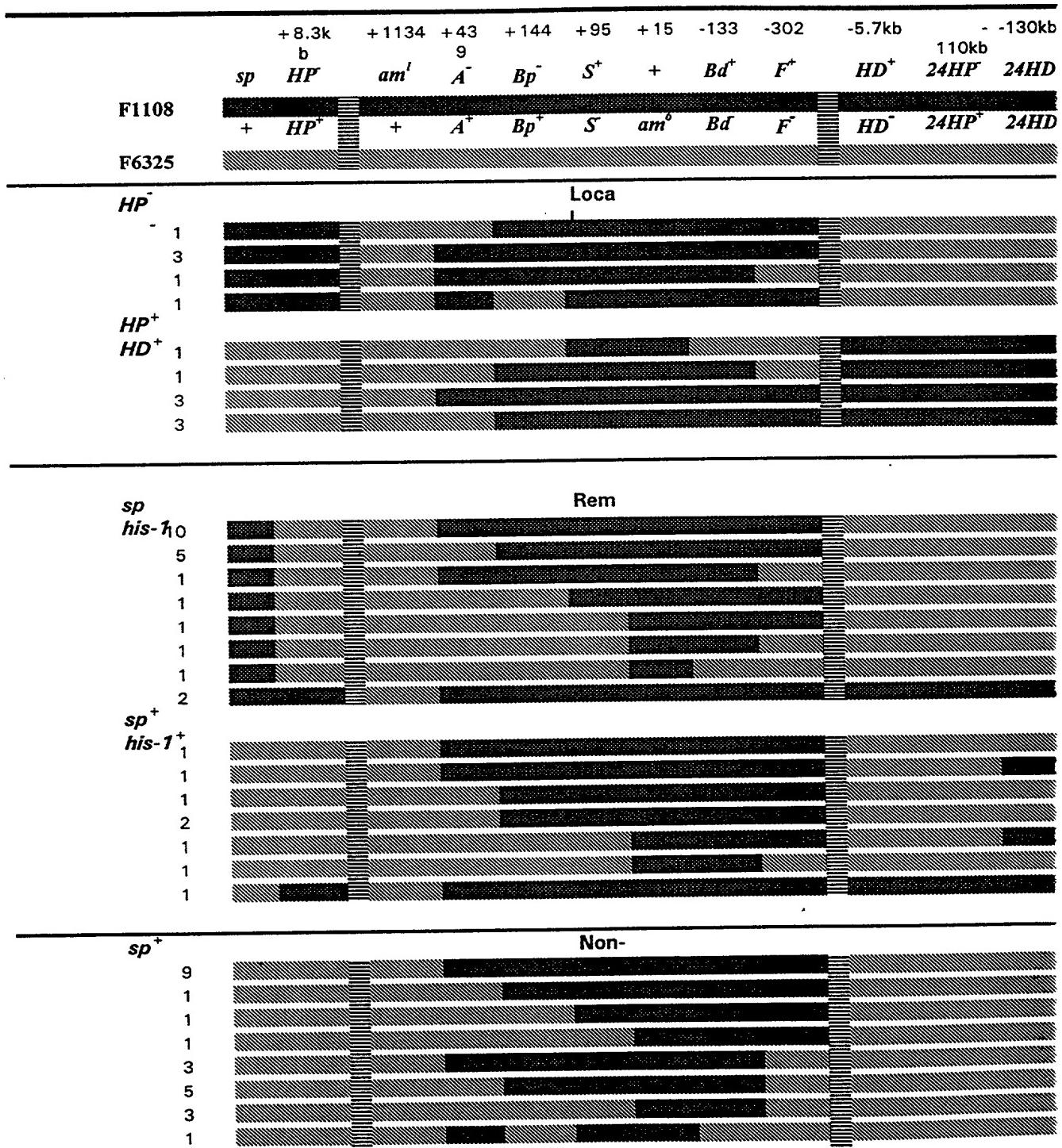


Fig 14  
(2 pages)



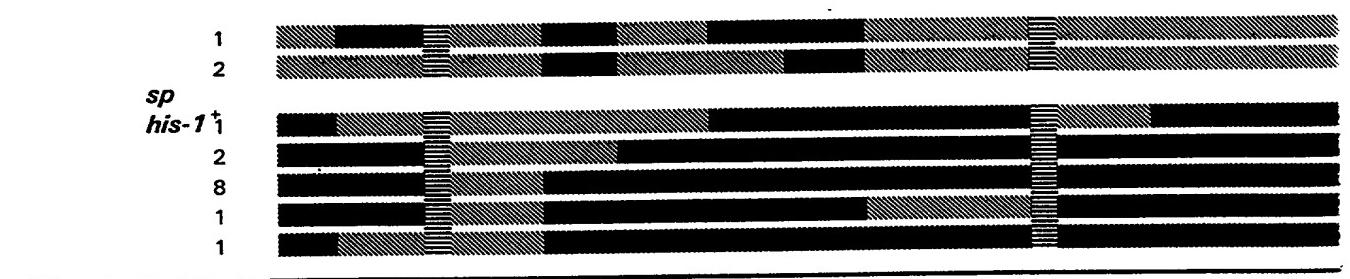
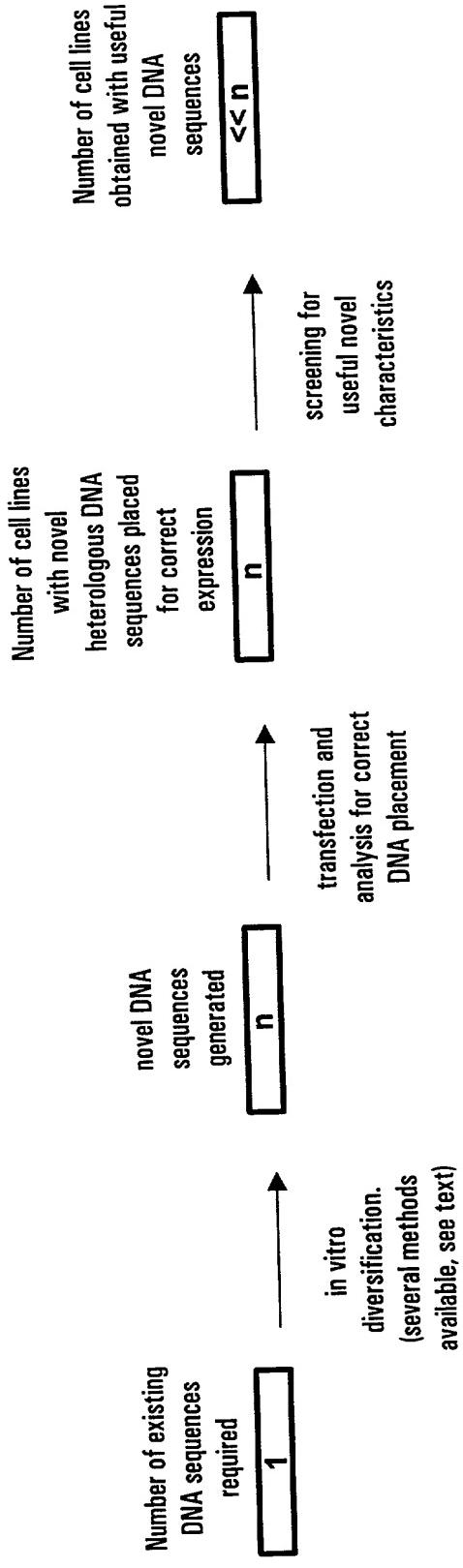


Fig. 14 (continued)

# FIG. 1A

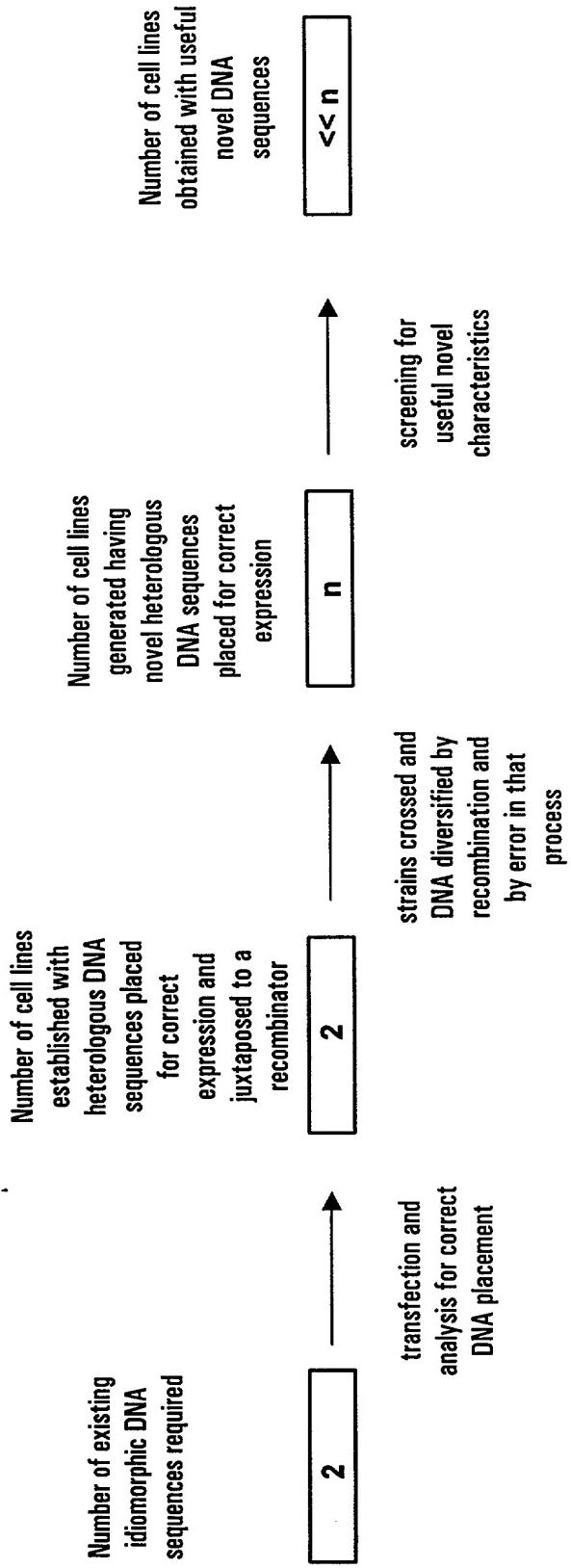
**Methods for the diversification of DNA sequences and testing for superior variants**

**existing protocols: Number of transfections needed to generate 1024 new variants: 1024**



## FIG. 1B

A protocol enabled by the present invention: **Number of transfections needed to generate 1024 new variants: 2**

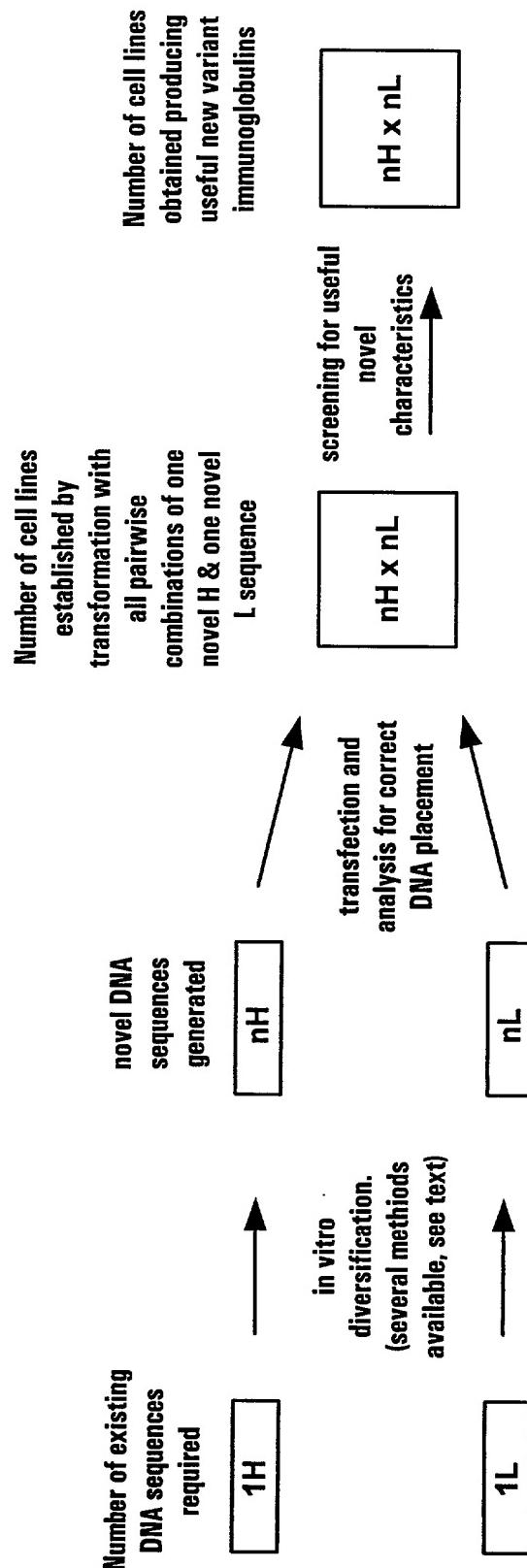


# FIG. 2A

**Methods for the diversification of DNA sequences coding subunits of heteropolymeric proteins and testing for superior variants.**

The example given for immunoglobulins is for illustrative purposes only and is not intended to limit application of the present invention to this specific heteromeric protein. H = heavy chain genes, L = light chain genes

**Existing protocol: Number of transfections needed to generate 1024 new combinations:** 2048



**FIG. 2B**

**Existing protocol using the heterokaryon technology of US Patent Serial No. 5,643,745**  
**Number of transfections needed to generate 1024 new combinations: 64**

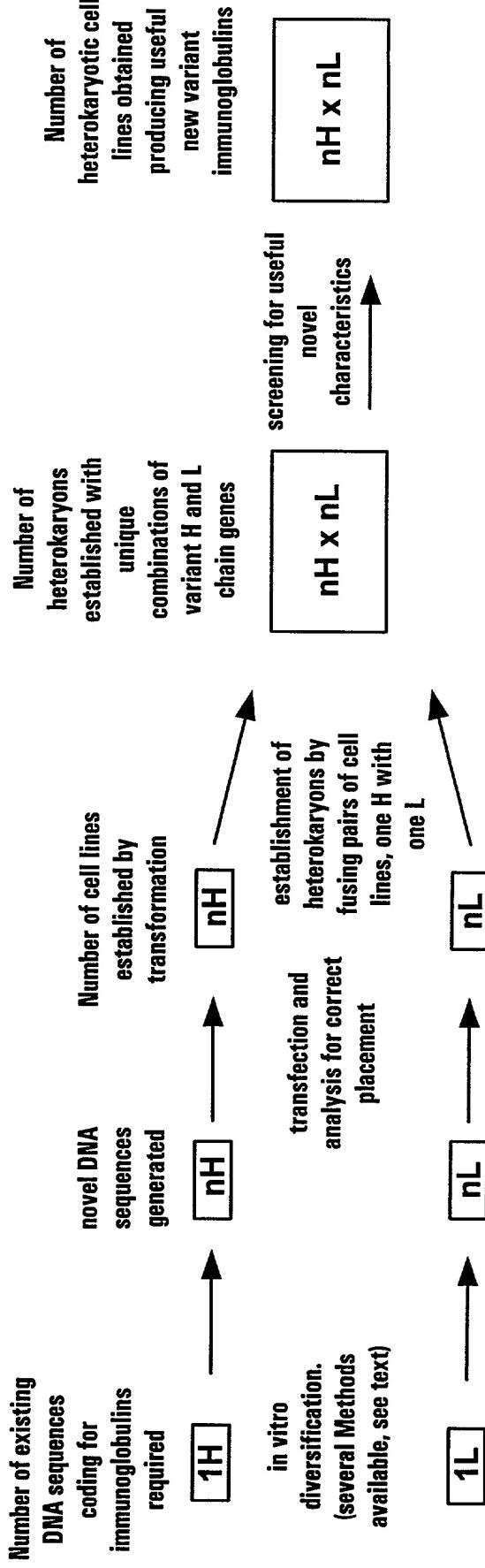
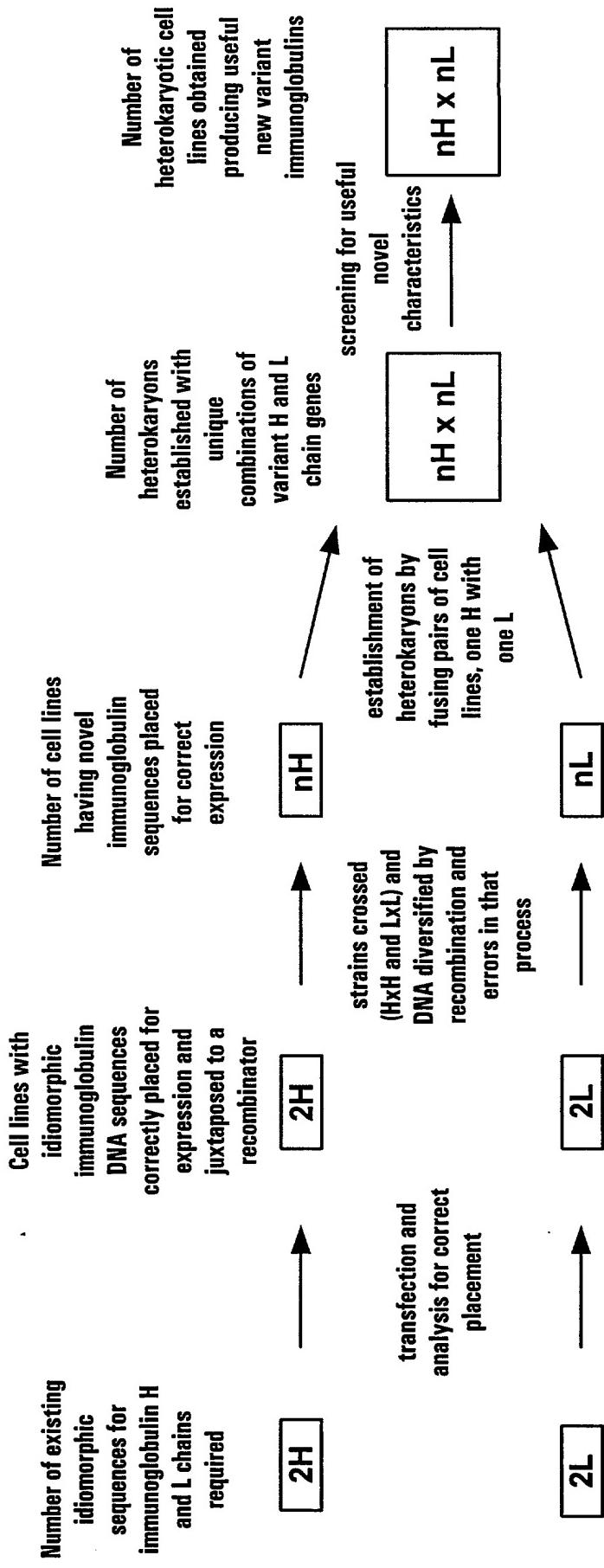


FIG. 2C

## A protocol enabled by the present invention

**Number of transfections needed to generate 1024 new combinations: 4**



## Fig. 3

The modified double strand break repair model for meiotic recombination. After H Sunet *al* Cell 64: 1155-1161, 1991

(a) A double strand break (DSB) is made in one DNA duplex. (b) A long 3' overhanging single strand tail is generated either side of the break by resection. (c) One 3' end invades a homologous duplex forming a D loop. (d) the D loop is enlarged by repair synthesis and anneals to the second 3' end (e) Repair synthesis occurs at the second 3' end and two intermolecular junctions (Holliday junctions) are formed. Resolution of the junctions by cutting inner and outer strands can give rise to non-crossover (f) and crossover (g) chromosomes. If there are base mismatches in the heteroduplex regions (duplex molecules with thick and thin lines) there will be gene conversion. If mismatch repair does not occur there will be post meiotic segregation of new sequence combinations.

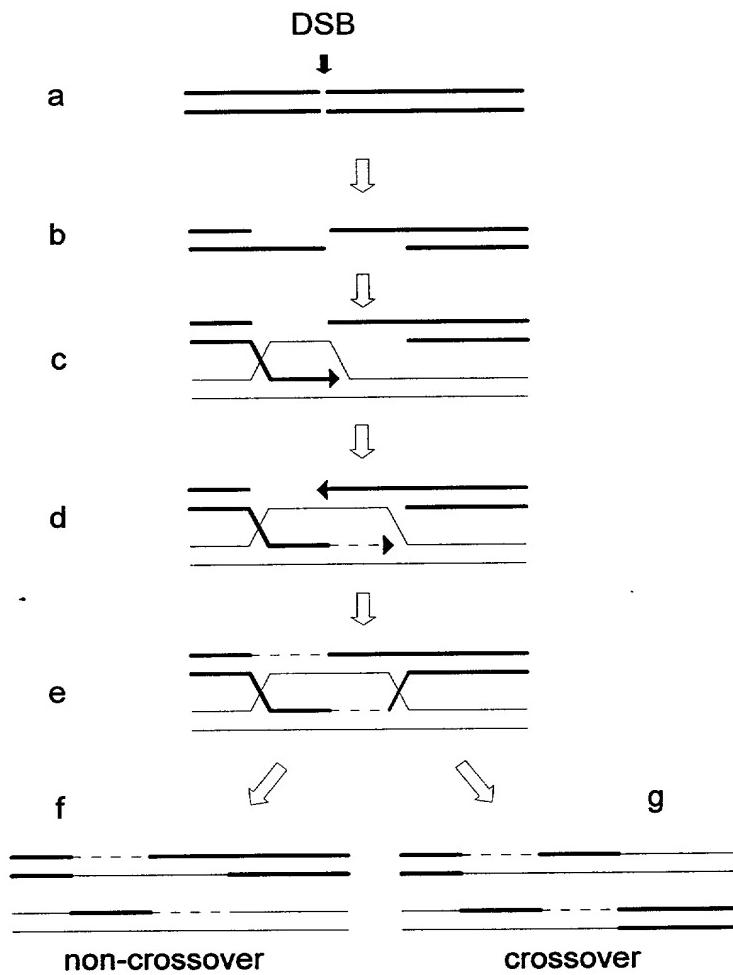
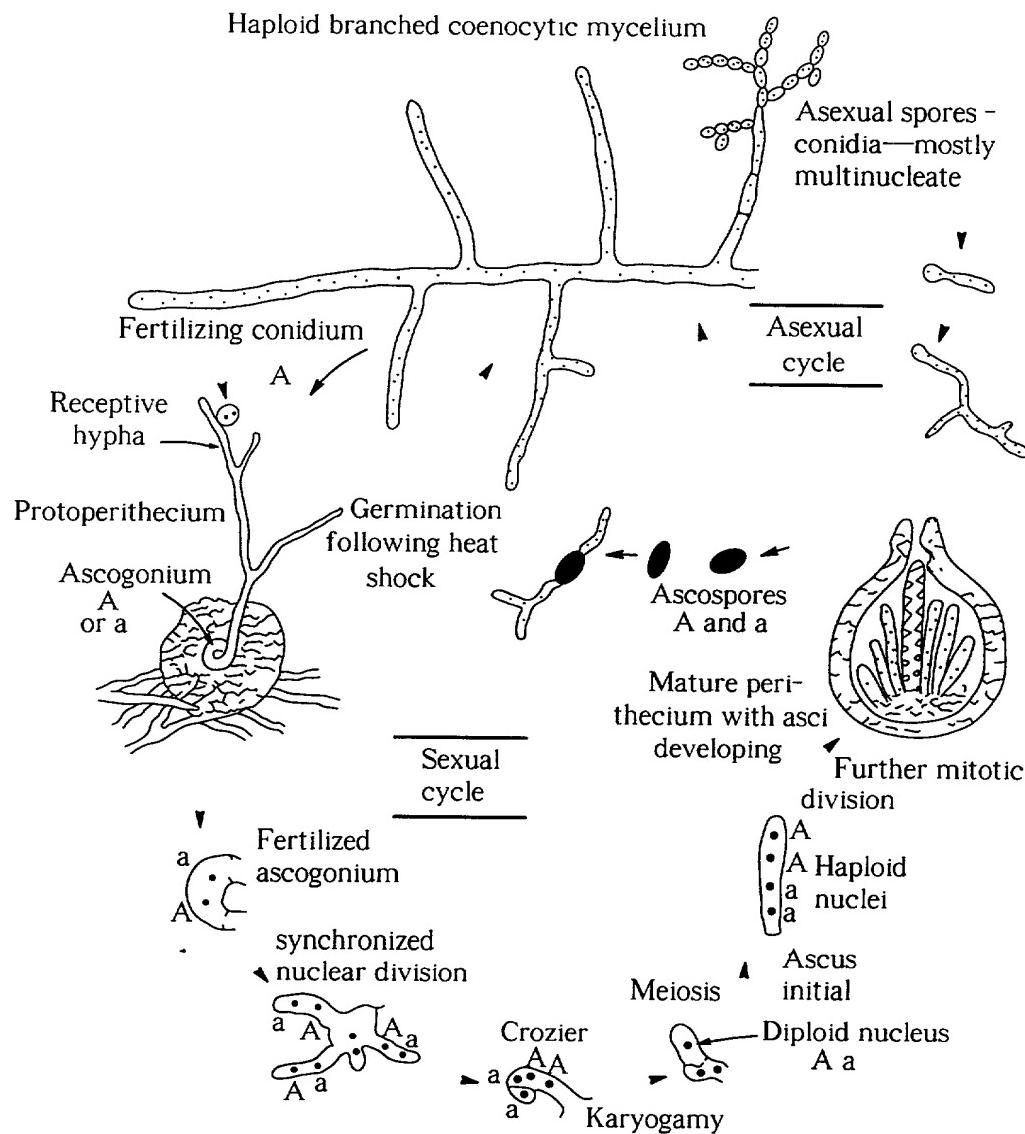
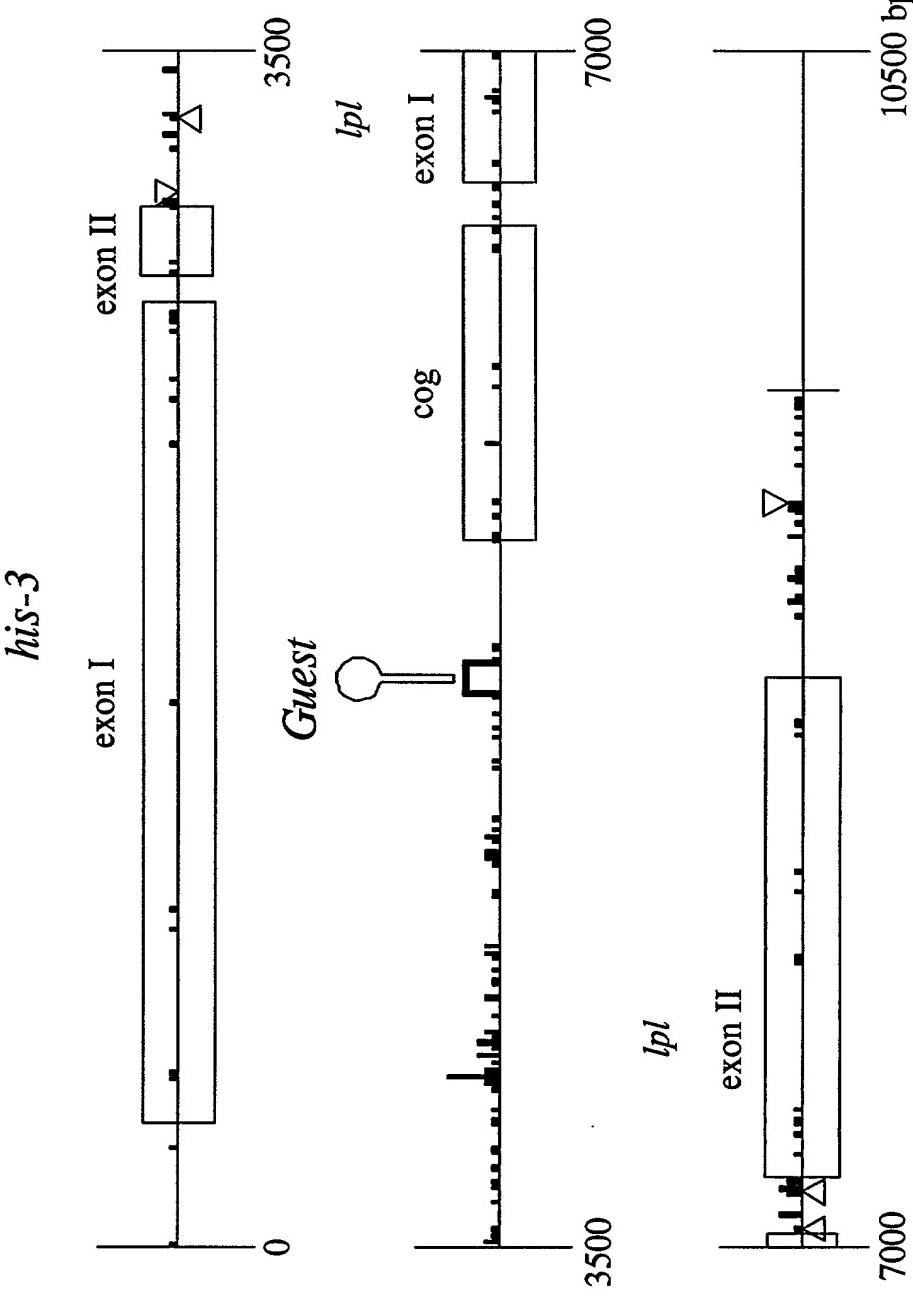


FIG. 4

Life cycle of *Neurospora crassa* after JRS Fincham (Genetics, Wright 1983). Microconidia having one nucleus are not shown but can be generated as described in the text. Perithecia and protoperithecia are shown in section.



**FIG. 5** Map of the *his-3*, *cog*, *ipl* region of Linkage Group I of *Neurospora crassa*. Vertical bars, triangles and hairpins show the location of sequence differences that distinguish the St Lawrence and Lindegren wild type strains. The corresponding full DNA sequences are given in Fig. 7 and Fig. 8. Vertical slashes indicate one to seven base substitutions per 10 base pairs. Triangles indicate short sequence insertions and the hairpin a 101 base pair inverted repeat transposon fragment present in St Lawrence.

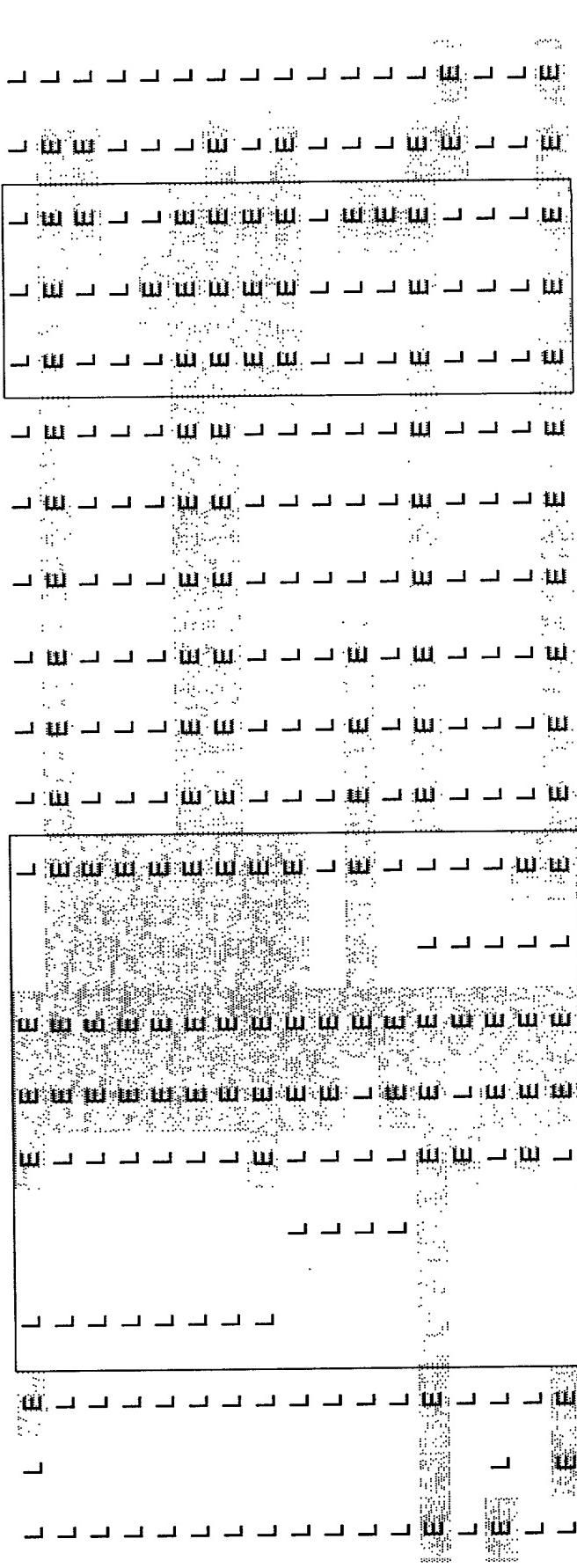


**FIG. 6** Discontinuity in the parental origin of DNA sequences in progeny from crosses between pairs of *his-3* alleles. In most cases this reflects discontinuity of conversion tracts, in some cases crossovers near the ends of conversion tracts. Markers are specific DNA sequence differences that distinguish the parents. These were all E (Emerson wild type origin) or all L (Lindegren wild type origin) in the parental strains. Recombinants carry both E and L markers. Marker position is given in base pairs from the first base of the first codon of the *his-3* gene. Each line of the table shows the parental origin of the markers inherited by one of the progeny.

marker location	P	H	P1	K1201	K504	L3	R1	K26	K874	R4	C4	C5	C6	C6/7	C7	C8	C1	C2	C3	C9	D
~66000	-384	115	179	563	1232	1502	1717	2318	3436	3705	4000	4304	4667	4821	5232	5495	6153	6507			

*his-3*

*cog*



# FIG. 7

Nucleotide sequence of the *his-3 cog<sup>L</sup> lpl* region of linkage group I in the Lindegren wild type strain of *Neurospora crassa*. This differs from that in the StLawrence strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the high frequency recombinator *cog<sup>L</sup>* which is active providing the cross in which meiosis occurs is homozygous *rec-2*.

1 GATCGCAACT GGAGATCACT CGCACCGTGC CGCAGAACAA GGGCGACGAG CCTCAGGGCA  
61 GTTTAGCCTG CCGTAACAGC ACAGACCATA GCTTATTTC ACCTGGCGG GCAGGGCGACG  
121 GCGGCACTGA CATCGGCAAG GCGGCATCAA GCAACCCCTC TGTTGCTTGC CAGCTGCCGG  
181 CCAACGTCAG CGGTACAAGG AGAAATCTGG AAGGAAAGAC TTCTGGCACC GACAGGATGG  
241 CACGCGGGAA AAGTTCCAA TGCATGAGAT GAGGGGCATT TGCATTGCCT CCCGTACAC  
301 TGCCCCGCAA CCCCACCCC ACCATAGCGT CTGTCGATAC ATGGAGCGCG AAGTCGAGAA  
361 ACCTGTAATT CCTGGTAACT TTCAGGTACA CAGTACGTAC TGATCCTGGT ATCAAACCTT  
421 GCCTGCCGAG TTTTCGACGG AAAGAGGTGT GAATTGTGAA AGAGTCATAC CAAATCACCC  
481 GATTTTCATA AAGCCCGAGT CTTTCTGTA CATAAGCGAC ACTCGAAGCG GCCCTCATCT  
541 TCATAGCCTG ATAGCTTGTA ATACTCCATC CTCGTATCTC ACTTGACCTT GAGTTCAACC  
601 CCACGTCAGA CTTCACCCGA CACATCGACG GATTGGGAA CAGCACAAATA CCTGAAAAGC  
661 GAGAAAACCA AACAGAGGAA AACACCATGG AGACAACACT TCCCCCTCCC TTCTCGTCG  
721 GTGTCAGTGT TCCTCCCCGA CTGAATGACA TCAAGGAGGG CCTCAGCCGG GAGGAAGTCT  
781 CGTGTCTTGG CTGCGTCTTC TTGAGGTCA AGCCCCAAGAC CCTTGAGAAA ATCGTGCAT  
841 TCCTCAAGCG TCACAATGTC GAATTGAGC CCTACTTCGA TGTAAACAGCC CTCGAGTCTA  
901 TCGATGATAT TATCACTCTT CTGGACGCC GCGCCCGCAA GGTGTTGTC AAGACCGAGC  
961 AGTTGGCCGA CCTCTCCGCA TATGGCTCCC GCGTTGCCCG CATTGTCACT GGAAGCAGCG  
1021 CTGCTTGCT TTCTCCGCC ACCGAGAGCG GCCTTTGCT CTCCGGCTTC GATCAGACTG  
1081 CCTCCGAGGC TGCACAGTT CTGGAGGAGG CCAGAGACAA GAAAATTACC CCCTTCTTCA  
1141 TCAAGCCCGT TCTGGGGCC GATCTCGAAC AGTTCATCCA GGTGCCGCC AAGGCTAACG  
1201 CCATCCCCAT CCTGCCATCC ACTGGCTTGA CAACAAAGAA GGACGGAGGCC GGAAAGCTTG  
1261 CCATCTCCAC CATCCTCTCG AGCGTCTGGA AGTCTGACCG TCCCGATGGT CTGCTCCCCA  
1321 CCGTTGTCGT TGATGAGCAC GACACTGCTC TGGGTCTGGT CTACAGCAGT GCCGAGAGTG  
1381 TGAACGAGGC CCTCAGGACA CAGACTGGTG TCTATCAGAG CCGGAAGCGC GGTCTCTGGT  
1441 ACAAGGGTGC TACTTCCGGA GACACTCAGG AGCTCGTCCG CATCTCGCTT GACTGCGATA  
1501 ACGATGCTCT CAAGTTGTC GTGAAGCAGA AGGGTCGTTT CTGCCACCTC GATCAGTCCG  
1561 GCTGTTGG TCAGCTAAA GGCCTTCCCA AGCTCGAGCA GACTTTGATT TCGAGGAAAC  
1621 AGTCTGCCCG CGAGGGCTCC TACACTGCC GTCTCTCTC CGATGAGAAG CTAGTCCGGG  
1681 CCAAGATCAT GGAGGAGGCT GAGGAGCTCT GCACCGCTCA GACCCCCCAG GAAATCGCCT  
1741 TTGAGGCTGC CGATCTCTC TACTTGCTC TTACCAAGGGC CGTTGCTGCC GGCAGTACTC  
1801 TTGCCGATAT CGAAAGGAGC CTTGACGCCA AGAGCTGGAA GGTCAAGCGC AGGACTGGAG  
1861 ATGCTAAGGG TAAGTGGGCT GAGAAGGAGG GCATCAAGCC TGCAGCGTCC GCTCCCGCTG  
1921 CCACTTCGGC CCCTGTCACC AAGGAGGCCG CCCAGGAGAC CACCCCTGAG AAGATCACCA  
1981 TGAGACGTTT CGACGCCCTCC AAGGCTCTA CCGAGGAGCT CGATGCTGCT CTCAAGCGTC  
2041 CTGCGCAAAA GTCGTCCGAT GCCATCTACA AGATCATTGT CCCCCATCATC GAGGACGTCC  
2101 GCAAGAACGG CGACAAGGCT GTTCTGTCGT ACACTCACAA GTTCAAGAAG GCTACCTCTC  
2161 TTACTAGCCC CGTCCTGAAG GCGCCCTTCC CCAAGGAGCT TATGCACTC CCTGAGGAGA  
2221 CCATTGCTGC CATCGACGTG TCCTTCGAGA ACATCGCAA GTTCCACGCC GCCCAGAAGG  
2281 AGGAGAAGCC CCTCCAGGTC GAGACCATGC CCGGTGTTGT CTGCAAGCCGT TTCTCTCGTC  
2341 CCATCGAGGC CGTCGGCTGC TACATCCCCG GCGGTACCGC CGTTCTCCCC AGCACTGCC  
2401 TTATGCTGGG TGTTCCCGCC ATGGTCGCCG GCTGCAACAA GATTGTGTT GCCTCTCC  
2461 CCCCGCGCCGA CGGAACCATC ACTCCCGAGA TTGTCACGT CGCTCACAAAG GTTGGGGCCG  
2521 AGTCCATCGT GCTTGCCGGC GGTGCCAGG CCGTAGCTGC CATGGCCTAC GGCACCGAGA  
2581 GCATCACCAA GGTGACAAG ATTCTCGGCC CCGGTAACCA GTTCACTC GCTGCCAAGA  
2641 TGTCGTCAG CAACGACACC AACGCTGCCG TTGGGATTGA CATGCCGCT GGCCCGTCCG  
2701 AGGTGCTGGT CATCGCTGAC AAGGACGCCA ACCCCGCCGT CGTTGCCTCG GATCTCCTGT  
2761 CCCAGGCTGA GCACGGCGTT GACAGTCAGG TCATCCTGAT CGCTATTAAC CTCGACGAGG

## FIG. 7 continued

2821 AGCATCTTCA GGCTATTGAG GACGGAGGTTG ACCGTCAGGC TATGGAGCTT CCTCGCGTCC  
2881 AGATTGTCCG TGGCTCCATC GCCCACTCGA TCACCGTGCA GGTCAAGACC GTCGAGGAGG  
2941 CCATGGAGCT CAGCAACAAG TACGCTCCTG AGCACTTGAT CCTCCAGATC AAGGAGGCCG  
3001 AGAAAGCTGT CGATCTTGTG ATGAACGCTG GTAGTGTCTT CATTGGCGCT TGGACTCCTG  
3061 AGTCCGTTGG CGATTACTCT GCTGGTGTAA ACCACTCGCT GCGTAAGTTA CATATCATAA  
3121 ATAGCCCCGC TTCACAGATT CTTCTGCTAA CGTCAAGACA CATAGCTACC TATGGTTTG  
3181 GCAAGCAGTA CTCTGGCGTC AATCTCGCT CGTCGTCAA GCACATTACC AGCTCCAAC  
3241 TGACTGCCGA GGGTCTCAAA AACGTCGGCC AGGCTGTCA GCAGTTGGCT AAGGTTGAGG  
3301 AGCTCGAGGC TCACAGAAGG GCGGTCAAGCA TCCGCTTGAGA GCACATGAGC AAGAGCAACT  
3361 AGACGGAAAT TCTTTTCGA AGTGCAAAA AAAACAAGAA CAAAAGGATG TAGTGGGTTG  
3421 ATGTATATCT GGGTCATTTT GGGCACATAG AGTAATGATA ACGAGTTTG GACATTGTAC  
3481 TGTTCTGTAC AGGCTGAAGA TCAGTACATG AATCTGTTGG TAAGTGTAGA GACCCAAACG  
3541 TCCCTTGAGT TTTCTCCCT GTTCCAGAGA GGTGCTCGTC CCTGGGTGTT TATTTTCATT  
3601 ATTACATCAA CCTTTTATT TATTTTATT TTTATTTAC TTTTTTTCC TTTTTTCAG  
3661 ATCATGCGTA CATGAACGGG GGAAGCACAG ACGATCGAAA CGTGGATGTC ACAATGTCGC  
3721 TGCAGTGTG AGGCTGTAC CTGCATTGCA TGAAGCGCCC ATCTCAATAT ACTTGCAGTC TTGCGCGTTG  
3781 CACGTGAACT TCCCAAACAA CCGAATAAAA GACGGCGAAA AATGAAGATA AAAAAAAACC  
3841 ATAATAAAAAA TCGGAGGGAG TGTGGGAAAT GGTTTCTTT AGCATTAGA CCCCATAGCC  
3901 GTGCACGCCA GGGTACAGAC AGGTTCATCG ATGTTGACAT TGACTGGGAC ACCAGGTCTA  
3961 TCTATTTCAT CTCCCTGTCC CTACCATACA TCGGGACATC GGACATCTCG CTGTACCCCC  
4021 CACACCCACA AAGTCTTATA AAAGCGCCAC ACCCGAGGAG GTTCGGTCGG CCCCACGAAC  
4081 TCCGTGCCTC CCTGCCTGTT TACAGGGACC GAACGCTGGA GAAGCTTAGT TTCCGTACAT  
4141 CGGGCCTACC CGAGCAGGAA AAGGGACAGC TCATAGGCAGA GGAGGGATTG GAAGATGGGG  
4201 ACATTTGGA TGATTGAGA GGAGGAACTA GGTACTGTAT CATGATAGTT CGGGGCAGCA  
4261 TCTTGGCTGG GACATTGTTA ATACCTCGAT ATGATGAAGT GGGAGGGAGT TTTTTCATGT  
4321 CTTGCCAAG TCCCACAAAT CTTTTTTTTT TTTTGTACCA ACACCCAAGA TTCTGGAGAAT  
4381 AGTGTAGGA TTCGCATTCA CAAGTGGAAAG TCTGAGGATC TTTTTATATC TTTGTCTTCC  
4441 GCGGACTGTT AACGATCCTA CAGCGAGCGA GCGAGCGTC GGATGCGCTG ATCTGTAGG  
4501 TGCAATATAAC GGCGCTTTC TCCGGTCGTG TAGTGTAAAGC TCTGTCGGCA TAGTAGTACA  
4561 CTAAAAAAAC CCTTGCAATT CATGATCTGC TTGCTATTCA TTCCGAGTTA TTTCACTGGT  
4621 CACATTTCGA GATTCACAGC CATCCATCCA TATGGAAAAA TCCATTCCCCA TGCTTCCCTCC  
4681 CCCCCACTAT GTATGTGACC ACACGCTGCT GTCAGAATGC CAACGGTCTC AGGTACCCCTC  
4741 GTCCGACTGT TTGGCATGGA GTTACATACA CTACTAGTGT AGCCCCGGGC CAAGCTACCC  
4801 CGTCAAATCT ATACATATCT ATAATGGGTT TCAGGTGTT CGTTCGCTGT CAATCAAGTT  
4861 TGAAACATCA CTGGGGCGT TGGACGGTGT ATTAGACCAT TGGCTCCCTC AGCTGGCGGC  
4921 TGGCGGTTG GGTGGCAAT AACGGGACTG GACTTGAGAG GGACGAGGAG AGTCGGTTGG  
4981 CTGCCTACAC TACACTACAA CGTTCACAC CTAACCGACG AGTCCCCTT TCCATTGTG  
5041 TGCCTTAACC ATCATCTAGG GATGTCAGGG TTTGGCCGA TCAGGGTATG TTTGGTTGAC  
5101 TGTTGTATG TCTGATGGG TACATATCAT GGTAGGTGTC TCGAGAACAG TAGAGTACTC  
5161 GGGCCTAGCG TTTGGATGAT TACGCGAGAT ATGAGTTGTA GGGCGCCATG CAGTTGCTTG  
5221 CCCATAAGCA GAAGTTGCTT TGGGATATAT TTCTCGTCTT TCAAAGGTCA CGAGGTCTG  
5281 GGACGAGCGG CATCGCCATC CAAAGGGTTG AACATGAGAA ACCGGAATGG CCTTGGCGTT  
5341 GAAATACAAA AAGTCAAGAA TAAAATCGCT TGAGGATAGG GACGTGGAAG CAAGCAAATA  
5401 TGGTAAGGGA GGTACTGCTA TGTAGGTGCT CAGCAAACG CCAATTCTT GGCCCCCAAG  
5461 CAGCAGTTG CTGTCAGTGC TGCTCGTGC AGCCTGGTA GTGGAACCTA AACTGCTAAC  
5521 ACAGCGCAAG TGCGCATGTA AAGATATTGT GGGAGGATCT GTATGGATGG ATGAGATTAC  
5581 TGCTTGGTGT TGGTTGCGAG GCACTGCGC TGTTAGGCTT TGCTGTGCC CGTTCGACGA  
5641 AGAAATACGC GGAACATATAA ATTGGATACC TAGACTTACT GCCTATGGGA GGTATCTACC  
5701 GACGTAGCCG ACGGATTCTA GCAACATCCC GACTTGCTT GTAGTGTACT ATGATAGCAG  
5761 CACAGTGGGG TGTTGCTCCT TGTGAGCATG GGCTCTTTT TTTTTTTCC CCCTTCCCTA  
5821 GGGCGTTGAC TGGACTTGCT CTATCGTCC CAAGGTAGGT GCCCGTCATC GATTTCCCA  
5881 AGCCGTCTCC CGCCAGATTG TCGTCATAGT GTCATGATGA CCTCGGTGCG TGGGGCTGCG  
5941 TGTTTACGGG GAGCTGGGAC CGCTAGGCCT CAGTGGTTGT GCCATTGAGC GTGGGGTGTGT  
6001 GGAGTAGCGG TAGAGGCCT TGGAAAGTTGT GCTAGCGGAA ACCCTGGAAT ATCTTGTACC

## FIG. 7 continued

6061 CTTCGATTCC TTCTCGGGCT GCCCATGTGC TGAGGTGATG CCGGGGATCT GGCGCCAATC  
6121 ATCCATTGAG GTTCCCGCAG CTTCCCGGTG CGCGCGCGG GCGCAGTTGC TCACAGGACA  
6181 CACCTAGACG CAGGGGCACA GGGGCACCGT TTGGTGTGCA ACTGGGTACC TGGTAGCTGT  
6241 AGCAAGCACT CCACCGTCTG TGCAATCCCC CAATCCACGG CAGGAACCTA GCACCGCCGC  
6301 GGCACCGAGT GAGCGAATCC ATCCGCATTG GATCCCAATT CTTGCCCTTG CCATCCCTCT  
6361 TTCTTCCCAC TTGGCGCAAC CAACACTTCC CTTGGTCTGG GTACTCGTGT TGATCTTCAC  
6421 TCTCTTTTT TCTTGGCGA CCGACTTTTT ATATCCGTCC TTGCTTCCCC CTGGCCGTTG  
6481 TCGTTCTTC TACAACCTACC TTCCGTTCAT TATCCCCTTT CTGGGTCGG TCGAGGACCC  
6541 AAAAACAGAA CAATTCCGGC TCTTCCAGGT GGCTGGGTG CGACTGTTA GCTCTTGACC  
6601 ACTAGCCGCT TACCTCTCT TGATGTTAT ATTTGGATAT CATTGAACTA CTCTTCTTG  
6661 AAACGGCAGA CGAACGGAAC AGTCCCTACG GTTTATTAGC GATATAACGTG GTACTGATAT  
6721 CCTGAGCAAG AAGAGGCAA TTATCAATTA TGCATCTCCC ATCGTCGCTG CTCATCGCAG  
6781 CTCCCTTGCT CGCCAATGTA TCGGCCGAAC CGATTAGGAT ACCCCAACGC GATGTTCTCC  
6841 GTGGTATCAA CATCACAGCA ACTTGCCGTT CGAGCACTAC CGAACATCGCC CAGCGGTGGA  
6901 TATGCCCTG CCGTTGAGA CTGTCCCAAG ACCAACCGA CGCTCCGGAA GGCGTGGAT  
6961 TTGTCGAACG AGGAGAAGAA CTGGTTGTCG ATCCGGAGGA AGAACACCAT CCAGCCCATG  
7021 AGGGACCTAC TGAAGAGGGC CAACATCACT GGGTCGATT CCGAAAACCTT CATGAATGAG  
7081 GCCGCCAAC ACGTCTCGCA ACTGCCAAT GTCGCCATTG CCATTCAGG AGGCGGCTAT  
7141 CGTGCCCTCA TGAACGGCGC CGGCTTCGTT GCTGCTGCGG ATAACCGGAT TCAAAATACC  
7201 ACGGGCGCAG GTGGTATTGG AGGCTTGTG CAGTCCAGCA CATATTTGTA TGTAAAACCA  
7261 TGCCTTCTTG TGGTTCTCT TATCTCGTT TCGAGTGTCA ACTGCGCCAG TTCGACGTTG  
7321 GGCGGCTGTG GACGACCTTG CTGGTGAACA TGTCTGGAC TCCATGCCCT TTTTTCTGTT  
7381 CCCTAAAATC CAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAATTCGAG  
7441 GACCGTGACT GTAAATTGCT AACGCAACTC TAGGGCCGGA CTTTCTGGTG GTGGCTGGCT  
7501 TGTCCGGAGT TTGTTCTCCA ACAACTTCAG TAGCATTGAG ACCCTGCTGA GCGAGAACAA  
7561 AGTCTGGGAC TTTGAGAACT CCATTTAA AGGACCAAG GAGGCTGGC TTAGTACTGT  
7621 CAACCGTATC CAGTACTGGT CGAAGTGGC AAAGGAAGTT GCGAAGAAGA AGGATGCTGG  
7681 CTTCGAGACA AGTATAACAG ACTACTGGGG CCGAGCATG AGTTACCAAC TGATCGGAGC  
7741 CGATATGGGC GGCCCGGCTT ACACCTTCTC CAGCATTGCC CAGACCGACA ACTTCCAGAA  
7801 GGCGAAACG CCGTCCCTA TTCTGGTAGC TGACGCCGC GGCCTGGAG ACACCATCAT  
7861 CTCCCTCAAT GCTACCAACT ACGAGTTCAA CCCGTTGAG ACGGGTAGCT GGGACCCGAC  
7921 CGTCTATGGC TTTGCGCCGA CCAAGTACCT CGGCCAAC TTCAGCAACG GCGTGTACCC  
7981 ATCGGGAGGC AAGTGCCTTG AGGGTCTCGA CCAAGCCGGC TTCGTATGG GCACCAGCAG  
8041 CACGCTCTTC AACCAGTTCC TTTTGCCAA CATCTCCAGC TACGACGGTG TTGCCAGACG  
8101 TGCTCATCGA GGCGTGACT TCTGCTCTCA AGGAAATCGG CGCAAGAGG ACGACGTCTC  
8161 CCAAATCATC CCTAATCCGT TCCTGGACTG GAACAACCCG ACCAACCCCA ACGCCGACAC  
8221 GCTCGAGCTC GACCTGGTCG ACGGCCGCAG AGATCTGCAG AATATTCCGC TCAACCCGCT  
8281 CACCAACCC GTGCGGCCG TCGACGTCACTT CTCGCTGTC GACTCGTCCG CCGACGTGAC  
8341 AAACGGCCC AATGGCACCG CCCTGCGCGC CACCTACGAG CGCACTTTCC GCTCTATTT  
8401 CAACGGGACA CTCTTCCCCT CGATCCCCGA CGACTGGACG TTATATAAACC TAGGCCTCAA  
8461 CAACCGCCCCC TCTTCTTCG GCTGCGATGT TAAGAACTTT ACCTGAACG CCAACCAAAA  
8521 GGTTCCCCCC TTAATCGTCT ATGTCCCCAA CGGCCCTAT ACCCGCCTGA GCAACGTGTC  
8581 CACCTTCGAT CGTCATACA CGATGTCTCA GCGCAACGAC ATCATCGCA ACGGATGGAA  
8641 CTCAGCCACG CAGGGAAACG GCACGCTGGA TTCGGAGTGG CCCACTTGCG TCGCCTGCG  
8701 GTTATCGAG AGGACCTTAG ATCGGTTGGG CAGGCAGACG CCAGCCGCGT GCAAGACTTG  
8761 CTTTGACAGG TATTGCTGGA ATGGCACAGT GAACTCCAA GATACGGGGG TTTACATGCC  
8821 TGAGTTCAAG ATTGGGGATG CGCATGCCCT GGACTCGGGT GCTGTTGCTA TCGGAAAGAT  
8881 GGTGAATGTC TGGTCGTCGG TTGTGGTGGG AGTTGTGGCG GCTACTTTGT TGTTGTAGGG  
8941 GTAGGGGAGA CGTGATGATA TTCCAGTCTG ATGAAGTTGA GACTGGACTG GAGATCGCCA  
9001 AGGATGCGGA GGGAAAGGA TCGCGTGGTGT TAATGTCATG ATGGATGAAG AGTCATGGAT  
9061 CATGGAACGA CGGGCGGGG ATATTGGATG ATGGATATAC CACACTGCAT GCATGCTCTA  
9121 TTGATAGTAT GCTTTGGCAT TTACGTTAA CAATCAATTG CTCCATCCTG ATGTTCTATC  
9181 TTTTCGACA ATGGATTGAT ACTACTCCTG TTGCTTCGCT CTGAGGTTG GAAGGACTTG  
9241 AGGTTGGAAG GACTTGAGGT TGTGTTGTTCT GAGGGAGGTT ATCGAAGTAT CATCTGTGCT

## FIG. 7 continued

9301 GATGCCGATT GATAGACTGT CCTCTTCTTC GAGGCAACGA ACGGTCGGAT GAGCCTCTTT  
9361 AATCATGATG CTCAGTGCCA CAAAAAGGCT CCAGCACAGC TGCCCACACC TTTCTTGCCCT  
9421 CGCCGTTCCCT TCCTTTTCTT TTTCCCCTGT TTCCCTTCTT CCTTCCATC TCATCCCCTA  
9481 CCAGAGTGCC CACCGGGTAT ATATATTACC TCCTTGCGCG TTCTCCTTTG ACCAATAAAT  
9541 CGCTTGGTCG AGTGGCGTAA CGGTTTACCG TCTACACTTA TCACTCAAAC CAAACCAAAC  
9601 CATCGAAGAA GTGACCTATC GGTCGAGGG AACGGTGATG TTCTTACGAC CAAGTTAAC  
9661 CAAAGAGCGT TCCACATCGT TGAACCGTCT CCTCCAGTTG GATCTGTTA ACTTCCGCAG  
9721 CGACTGAAGA AGGTATCACT TTTTTTTGG TTCCAAAAAA AAAAAAAA ATTAC

## FIG. 8

Nucleotide sequence of the *his-3 cog<sup>E</sup> lpl* region of linkage group I in the StLawrence wild type strain of *Neurospora crassa*. This differs from that in the Lindegren strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the weak recombinator *cog<sup>E</sup>* and also the remnant of a transposable element *Guest* within the replaceable sequence 3' of *his-3*. StLawrence strains carry *rec-2<sup>+</sup>* which prevents the initiation of recombination at *cog*.

```

1 ACCGGGAATC GTAGCGGGCG CTAAGGCCAA GCGCGGGCAC GGGTCACTGA CCCAATGCAG
61 CGCATTCGGT CAGCAACTGA AGTGGATGTA CAAGTACATA GTAGTAGATC GCAACTGGAG
121 ATCACTCGCA CCGTGCCGCA GAACAAGGGC GACGAGCCTC AGGGCAGTTT AGCCTGCCGT
181 AACAGCACAG ACCATAGCTT ATTTTCACCT GGGCGGGCGG GCGACGGCGG CACTGACATC
241 GGCAAGGCAG CATCAAGCAA CCCCTCTGTT GCTTGCCAGC TGCCGGCCAA CGTCAGCGGT
301 ACAAGGAGAA ATCTGGAAGG AAAGACTTCT GGCACCGACA GGATGGCACG CGGGAAAAGT
361 TCCCAATGCA TGAGATGAGG GGCATTTGCA TTGCCTCCCG TCACCCAGTG CGAACCCCAA
421 CCCCCACCATA CGCTCTGTGCG ATACATGGAG CGCGAAGTCG AGAAACCTGT AATTCTGGT
481 AACTTTCAGG TACACAGTAC GTACTGATCC TGGTATCAAA CCTTGCCTGC CGAGTTTCG
541 ACGGAAAGAG GTGTGAATTG TGAAAGAGTC ATACCAAATC ACCCGATTTC CATAAAGCCC
601 GAGTCTTTC TGTACATAAG CGACACTCGA AGCgggcctc ATCTTCATAG CCTGATAGCT
661 TGTAACTAC TACCTCGTA TCTCACTTGA CCTTGAGTTC AACCCACGT CAAACTTCAC
721 CCGACACATC GACGGATTGG GGAACAGCAC AATACCTGAA AAGCGAGAAA ACCAACAGA
781 GGAAAACACC ATGGAGACAA CACTTCCCTT CCCCTCCTC GTCGGTGTCA GTGTTCTCC
841 CGGACTGAAT GACATCAAGG AGGGCCTCAG CGGGGAGGAA GTCTCGTGT TTGGCTGCGT
901 CTTCTTCGAG GTCAAGCCCC AGACCCCTTGA GAAAATCCTG CGATTCTCA AGCGTCACAA
961 TGTGAATTG GAGCCCTACT TCGATGTAAC AGCCCTCGAG TCTATCGATG ATATTATCAC
1021 TCTTCTGGAC GCCGGCGCCC GCAAGGTGTT TGTCAAGACC GAGCAGTTGG CCGACCTCTC
1081 CGCATATGGC TCCCGCGTTG CCCCCATTGT CACTGGAAGC AGCGCTGCTT TGCTTCTCC
1141 CGCCACCGAG AGCGGCCCTT TGCTCTCCGG CCTCGATCAG ACTGCCTCCG AGGCTGCACA
1201 GTTTCTGGAG GAGGCCAGAG ACAAGAAAAT TACCCCCCTC TTCATCAAGC CCGTTCTGG
1261 GGCGATCTC GAACAGTTCA TCCAGGTGCG CGCCAAGGCT AACGCCATCC CCATCTGCC
1321 ATCCACTGGC TTGACAACAA AGAAGGAGCA GGGCGGCAAG CCTGCCATCT CCACCATCCT
1381 CTCGAGCGTC TGGAGTCTG ACCGTCCCCA TGGTCTTCTC CCCACCGTTG TCGTTGATGA
1441 GCACGACACT GCTCTGGTC TGGTCTACAG CAGTGGCAGAG AGTGTGAACG AGGCCCTCAG
1501 GACACAGACT GGTGTCTATC AGAGCCGGAA GCGCGGTCTC TGGTACAAGG GTGCTACTTC
1561 CGGAGACACT CAGGAGCTCG TCCGCATCTC GCTGACTGC GATAACGATG CTCTCAAGTT
1621 TGTGTTGAAG CAGAAGGGTC GTTCTGCCA CCTCGATCAG TCCGGCTGCT TTGGTCAAGCT
1681 CAAAGGCCCTT CCCAAGCTCG AGCAGACTTT GATTGAGG AAACAGTCTG CCCCCGAGGG
1741 CTCCTACACT GCCCGTCTCT TCTCCGATGA GAAGCTAGTC CGGGCCAAGA TCATGGAGGA
1801 GGCTGAGGAG CTCTGCACCG CTCAGACCCC CCAGGAAATC GCCTTGAGG CTGCCATCT
1861 CTTCTACTTT GCTCTTACCA GGGCCGTTGC TGCCGGCGTT ACTCTGCCC ATATCGAAAG
1921 GAGCCTGAC GCCAAGAGCT GGAAGGTCAA GCGCAGGACT GGAGATGCTA AGGGTAAGTG
1981 GGCTGAGAAG GAGGGCATCA AGCCTGCGGC GTCCGCTCTC GCTGCCACTT CGGCCCTGT
2041 CACCAAGGAG GCCGGCCAGG AGACCACCCC TGAGAAGATC ACCATGAGAC GTTCTGACGC
2101 CTCCAAGGTC TCTACCGAGG AGCTCGATGC TGCTCTCAAG CGTCTGCGC AAAAGTCGTC
2161 CGATGCCATC TACAAGATCA TTGTCCCCAT CATCGAGGAC GTCCGCAAGA ACGGCGACAA
2221 GGCTGTTCTG TCGTACACTC ACAAGTTCGA GAAGGCTACC TCTCTTACTA GCCCCGTCTC
2281 GAAGGGCCCTT TTCCCCAAGG AGCTTATGCA GCTCCCTGAG GAGACCATGG CTGCCATCGA
2341 CGTGTCTTC GAGAACATCC GCAAGTTCCA CGCCGCCAG AAGGAGGAGA AGCCCCCTCCA
2401 GGTCGAGACC ATGCCCGGTG TTGTCTGCA CGTCTCTCT CGTCCCATCG AGGCCGTCTG
2461 CTGCTACATC CCCGGCGGTG CGCGCGTTCT CCCACACT GCCCTTATGC TGGGTGTTCC
2521 CGCCATGGTC GCCGGCTGCA ACAAGATTGT GTTCCCTCT CCTCCCCGGC CCGACGGAAC
2581 CATCACTCCC GAGATTGTCC ACGTCGCTCA CAAGGTTGGG GCGAGTCCA TCGTGTGTTGC
2641 CGGCGGTGCC CAGGGCGTAG CTGCCATGGC CTACGGCACC GAGAGCATCA CCAAGGTCGA
2701 CAAGATTCTC GGCCCCGGTA ACCAGTTCGT CACTGCTGCC AAGATGTTCG TCAGCAACGA

```

## FIG. 8 continued

2761 CACCAACGCT GCCGTTGGTA TTGACATGCC CGCTGGCCCG TCCGAGGTGC TGGTCATCGC  
2821 TGACAAGGAC GCCAACCCCCG CGTTCGTTGC CTCGGATCTC CTGTCCTCAGG CTGAGCACGG  
2881 CGTTGACAGT CAGGTCATCC TGATCGCTAT TGACCTCGAC GAGGAGCATT TTCAGGCTAT  
2941 TGAGGACGAG GTTCACCGTC AGGCTACGGA GCTTCCTCGC GTCCAGATTG TCCGTGGCTC  
3001 CATCGCCCAC TCGATCACCG TGCAAGTCAA GACCCTGAG GAGGCCATGG AGCTCAGCAA  
3061 CAAGTACGCT CCTGAGCACT TGATCCTCCA GATCAAGGAG GCCGAGAAGG CTGTCGATCT  
3121 TGTCAATGAAAC GCCGGTAGTG TCTTCATTGG CGCCTGGACT CCTGAGTCCG TTGGCGATTA  
3181 CTCTGCTGGT GTTAACCACT CGCTGCGTAA GTTACATATC ATAAATAGCC CCGCTTCACA  
3241 GATTCTTCTG CTAACGTCAA GACACATAGC TACCTATGGC TTTGGCAAGC AGTACTCTGG  
3301 CGTCAATTTC GCCTCGTTCG TCAAGCACAT TACCAAGCTCC AACTTGACTG CCGAGGGTCT  
3361 CAAAAAACGTC GGCCAGGCTG TCATGCAGTT GGCTAAGGTT GAGGAGCTCG AGGCTCACAG  
3421 AAGGGCGGTC AGCATCCGTC TTGAGCACAT GAGCAAGAGC AACTAAACGG AAATTCTTT  
3481 CGAAGTAGCA AAAAAAAA AAAAAAACAA GAACAAAAGG ATGTAGTGGG TTGATGTATA  
3541 TCTGGGTCA TTTGGGCACA TAGAGTAATG ATAACGAGTT TTGGACATTG TACTGTTCTG  
3601 TACAGGCTGA AGATCAGTAC ATGAATCTGT TGGTAAGTGT GGAGACCCAA ACGTCCCTTG  
3661 AGTTTTTCTC CCTATTCCAG AGGTGCTCGT CCCTGGGTGT TTATTTTCA TATTACATCA  
3721 ACCTTTTTT TTTTTTTTT TTTTCAGAT CATCGTACA TGAACGGGG AAGCACAGAC  
3781 GATCGAAACG TGGATGTCAC AATGTCGCTG CAGTGATGCT GCATTGCATG AAGCGCCCAT  
3841 CTCAATATAC TTGCAGTCTT GCACGTTGCA TGTGAACCTC CAAACAAACC GAATAAAAGA  
3901 CGGCGAAAAAA TGAAGATAAA AAAAAACCAT AAAAAAAATC AGAGGGAGTG TGGGAATGG  
3961 TGTCTTTAG CATTCAAGACC CCATAGCCGT GCACGCCCG GTACAGACAG GTTCATCGAT  
4021 GTTGACATTG ACTGGGACAC CAGGTCTATC TATTTTATCT CCTGTCCTCT ACCATACATC  
4081 GGGACATCGG ACATCTGCT GTACCCCCCA CACCCACAAA GCCTTATAAA AGCGCCACAC  
4141 CCGAGGAGGT TCGGTCGGCC CCACGAACTC TGTGCCTCCC TGCCTGTTA CAGGGACCGA  
4201 ACGCTGGAGA ATCTTACTAG TTTCCTGACA TCCGGCCTAC CCGAGCAGGA AAAGGGACAG  
4261 CTCATAGGCG AGGAGGGATT TGAAGATGGG AACATTTGG GTGATTGAG AGGAGGAAC  
4321 AGGTACTGCA TCATGATAGT TCGGGGCAGC ATCTTGGCTG GGACATTGTT AATACCTCGA  
4381 TATGATGAAG TAGGAGGGAG TTTTGCGTG TCTTGCCGAA GTCCAGAGAT CTGTTTATT  
4441 TTATTTTTA TGGATGTAGT GTATCAACAC CCAAGATTG GAGAATAGTA CTAGGATTG  
4501 CATTTACAAG TGGAACTCTT GAGAATCGTT GTATATCCTT GTCTTCCTCG GAATGTTAAC  
4561 AATCCTACAG CGAGCGAGCG AGCGGTCGGA TGCCTGATC TGATAGGCGC AATATACGGC  
4621 CGCTTCTCC GGTCGTGAG TGTAAGCTCT GTGGGCATAG TACACTAAAA AAACCCCTTG  
4681 ATTTCATGAT CTGCCTGCTA TTCATTCCGA GCTATTTCA TGGTCACATT TCGAGGAAGA  
4741 AAGAAAGCAA CTAAGATTCA CAGCCATCCA TCCATCCATA TGGAAAGAATA ATCCATTCCC  
4801 ATGTTCCCTC CCCCCCACTA TGTATGTGAC CACACGCTGC TGTCAAGATG CCAACGGTCT  
4861 CAGGTACCCCT CGTCCGACTG TTTGGCATGG AGTTACATAC ACTACTAGTG TAGCCCCGGG  
4921 CCAAGCTACC CCGTCAAATC TATACATATC TATAACGGGT TTCAGGGGTT TCGTTCGCTG  
4981 TCAATCAAGT TTGAAACATC ACTGGGGCCG TTGGACGGTG TATTAGACCA TTGGCTCCCT  
5041 CAGCTGTTG GCGGCTGGC GGCTGGGTCA AACGGAATA ACGGGACTCG AGAGGGACGA  
5101 GGAGAGTCGG TTGGCTGGCT GCAATACAAG CGTCCCACC TAACCAACGA GTCCCGTTT  
5161 CCATTTGTGT GCCTAACCAT CATCTAGGGG TGTCAGGGTT TGGCCGGATC AGGGTATGTT  
5221 TGGTTGACTG TTGTCATGTC TGATTGGGT CATATTATGG TAGGTGTCTC GAGAACAGTA  
5281 GAGTAACCGG GCCTAGCGTT TGGATGATTA CGCGAGATAT GAGTTGTGGG CCGCCATGCA  
5341 GTTGCTTGTG CATAAGCAGA AGTTGCTTG GGATATATT CTCGCTTTC AAAGGTACAG  
5401 AGGTCTGGG ACGAACGGCA TCGCCATCCA AAGGGTTGAA CATGAGAAAC CTGAATGGCC  
5461 TTTGCGTTGA AATACAAAAA GTCAAGAACAA AATCGCTG AGGATAGGGA CGTGGAAAGCA  
5521 AGCAAATATG GTAAGAGAGG TATACATCAA CCCTGGTCA ATTGTTAGCG TGGTTCTTCC  
5581 TCCACGTCT CGTTCATGAC GGTAAACAGT ACCAGGCTAA CAATTAAACC AGGGTTGATG  
5641 TGTACTGATA TGTAGGTGCT CAGCAAACG CCAATTCTT TGGCCCCAAG CAGCAGTTG  
5701 CTGTCAGTGC TGCTCGTGC AGCCTGGTA GTGGAACCTA AACTGCTAAC ACAGCGCAAG  
5761 TGCGCATGTA AAGATATTGT GGGAGGATCT GTATGGATGG ATGAGATTAC TGCTGGTGT  
5821 TGGTTGCGAG GCACTGCGGC TGTTAGGCTT TGCTGTGCC CGTCGACGA AGAAATACGC  
5881 GGAACATATAA ATTGGATACC TAGACTTACT GCCTATGGGA GGTATCTACC GACGTAGCCG  
5941 ACGGATTCTA GCAACATCCC GACTTGCTT GTAGTGTACT ATGATAGCAG CACAGTGTG

## FIG. 8 continued

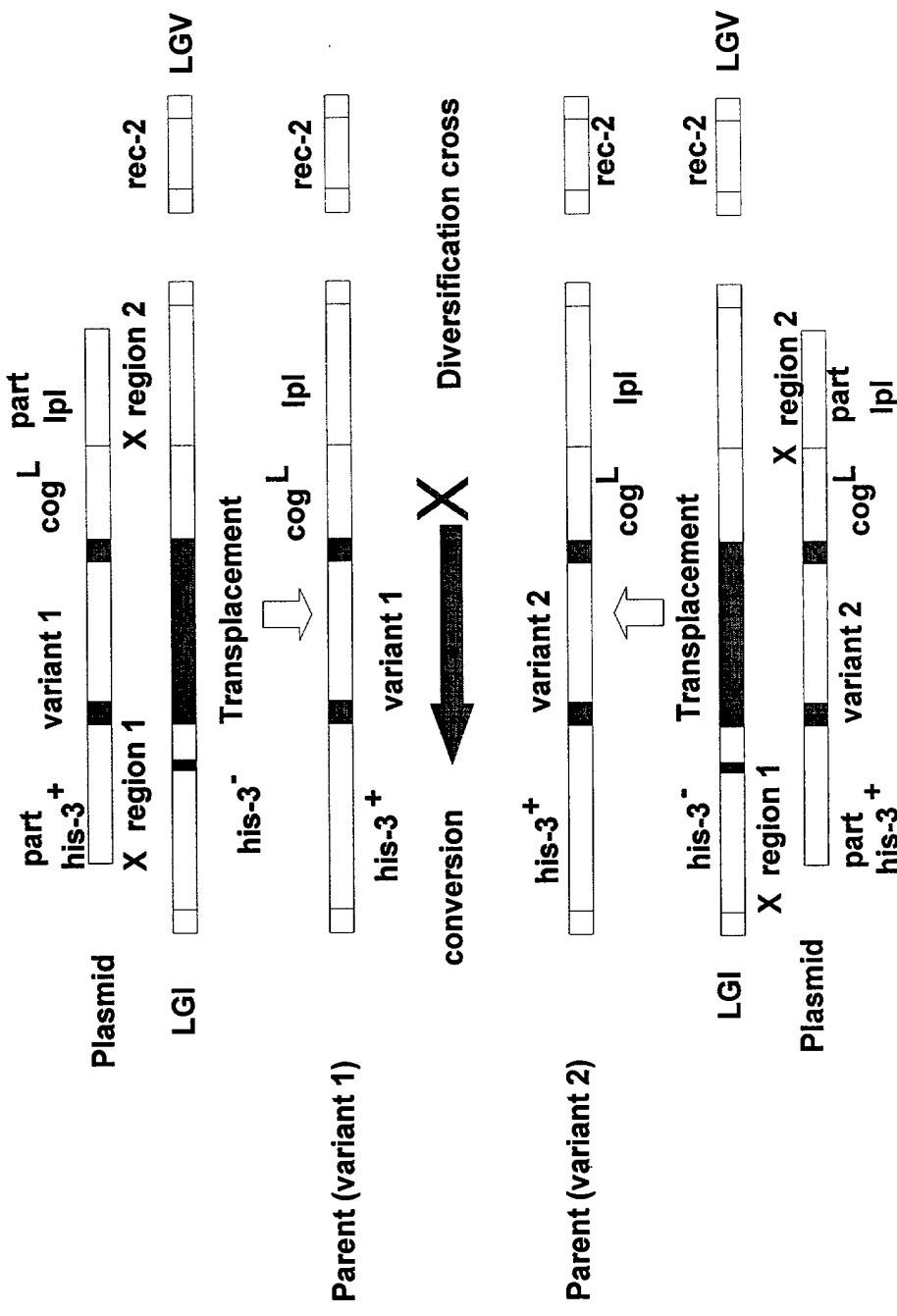
6001 CTCCCTTGTGA GAATGGGCTC TTTTTTTTT TCCCCCTTCC CTAGGGCGTT GACTGGACTT  
6061 GCTCTATTGT TCCCAAGGTA GGTGCCGTC ATCGATTTTC CCAAGTCTCC CGCCAGATTG  
6121 TCGTCATAGT GTCATGATGA CCTCGGTGCG TGCGGTGCG TGTTACGGG GAGCTGGAC  
6181 CGCTAGGCCT CAGTGGTTGT GCCATTCAAG GTGGGTGTGT GGAGTAGCGG TAGAGGCCT  
6241 TGGAAAGTTGT GCTAGCGGAA ACCCTGGAAT ATCTTCTACC CTCGATTCCCT TCTCGGGCTG  
6301 CCCATGTGCT GAGGTGATGC CGGGGATCTG GCGCCAATCA TCCATTGAGG TTCCCGCAGC  
6361 TTCCCGGTGC CGCGCGCGG CGCAGTTGCT CACAGGACAC ACCTAGACGC AGGGGCACAG  
6421 GGGCACCGTT TGGTGTGCAA CTGGGTACCT AGCTGTAGCA AGCACTCCAC CGTCTGTGCA  
6481 ATCCCCCAAT CCACGGCAGG AACTTCGCAC CGCCGGGCA CCGAGTGAGC GAATCCATCC  
6541 GCATTGGATC CCAATTCTTGC CCCTTGCAT CCTTCTTCT TCCCACCTGG CGCAACCAAC  
6601 ACTTCCCTTG GTCTGGGTAC TCGTGTGAT CTTCACTCTC TTTTTTCTT GGGCGACCGA  
6661 CTTTTTATAT CCGTCTTGC TTCCCCCTGG CGTTGTCGT TCTTCTACA ACTACCTCC  
6721 GTTCATTATTC CCCTTCTTG GTTCGGTCA GGACCCAAA ACAGAACAAAT TCCGGCTCTT  
6781 CCAGGTGGCT TGGGTGCGAC TGTTAGCTC TTGACCACTA GCGCCTTACCG TTCTCTTGAT  
6841 GTTTTTATTT GGATATCATT AAACACTCT TTCTGAAAC GGCAGACGAA CGGAACAGTT  
6901 CCTACGGTAT ATTAGCGATA TACGTTGTAC TGATATTCTG AGCAAGAAGA GGCAAATTAT  
6961 CAATTATGCA TCTCCCTTCG TCGCTGCTCA TCGCAGCTCC CTTGCTCGCC AATGTATCGG  
7021 CCGAACCCAT TAGGATACCC CAACCGCATG TTCTCCGTGG TATCAACATC ACAGCAACCTT  
7081 GCGTTCGAG CACTACCGGA TTCGCCAGC GGTGGATATG CCCCTGCCGT TGTAGACTGT  
7141 CCCAAGACCA AGCCGACGCT CCGGAAGGCC GTGGATTGT CGAACGAGGA GAAGAACTGG  
7201 TTGTCGATCC GGAGGAAGAA CACCATCCAG CCCATGAGGG ACCTCCTGAA GAGGGCCAAC  
7261 ATCACTGGGT TCGATTCCGA GACATTATG AATGAGGCCG CCAACAAACAT CTCGCAACTG  
7321 CCCAATGTGCA CCATTGCCAT TTCAGGAGGC GGCTATCGT CCCTCATGAA CGCGGCCGGC  
7381 TTCGTTGCTG CTGCGATAA CGAATTCAA AATACCACGG GCGCAGGTGG TATTGGAGGC  
7441 TTGTTGCAGT CCAGCACATA TTTGTATGTA AAGTGGTCT TCTTATCTCG TTTTCGAGTG  
7501 TCAACTGCGC CAGTCAGAG TTGGCGGGT GTGGACGACC TTGCTGGTGA ACATGTCTTG  
7561 GACTCCATGC CCCTTCTTCG TTTCTCAAA TCAAGAAGTC GAGGACCGTG ACCGTAATC  
7621 GCTAACGCAA CTCTAGGGCC GGACTTCTG GTGGTGGCTG CTTGTCGCG AGTTTGTCT  
7681 CCAACAACTT CAGCAGCATT GAGACCCCTGC TGAGCAGGAA CAAAGTCTGG GACTTGTAGA  
7741 ACTCCATCTT TAAAGGGCCC AAGGAGGCTG GCCTTAGTAC TGTCAACCGC ATTCACTACT  
7801 GGTCCGAAGT GGCAAAAGGAA GTTGCAGAAGA AGAAGGATGC TGGCTTCGAG ACAAGTATAA  
7861 CAGACTACTG GGGCCGAGCA TTGAGTTACC AACTGATCGG AGCCGATATG GGCAGGCCGG  
7921 CTTACACCTT CTCCAGCATT GCCCAGACCG ACAACTTCCA GAAGGCCGAA ACGCCGTTCC  
7981 CTATTCTGGT AGCTGACGGC CGCGCGCTG GAGACACCAT CATCTCCCTC AATGCTACCA  
8041 ACTACGAGTT CAACCCGTTG GAGACGGGTA GCTGGGACCC GACCGTCTAT GGCTTGC  
8101 CGACCAAGTA CCTCGCGCC AACTTCAGCA ACGGCGTGT CCCATCGGGG GGCAAGTGC  
8161 TTGAGGGTCT CGACCAAGCC GGCTTCGTCA TGGGCACCAG CAGCACGCTC TTCAACCACT  
8221 TCCTTTGGC CAACATCTCC AGCTACGACG GTGTTGCCG ACGTGCTCAT CGAACCGTG  
8281 ACTTCTGTCC TCAAGGAAAT CGCGCCAAG AGGACGACGT CTCCCAAATC ATCCCTAAC  
8341 CGTTCTGGA CTGGAACAAC CGGACCAACC CCAACGCCA CACGCTCGAG CTCGACCTGG  
8401 TCGACGGCGG CGAAGATCTG CAGAATATTC CGCTCAACCC GCTCACCCAA CCCGTGCG  
8461 CGTGGACGT CATCTCGCT GTGACTCGT CCGCCGACGT GACAAACTGG CCCAATGGCA  
8521 CGCCCTGCG AGCCACCTAC GAGCGCACTT TCGGCTCTAT TTCAACGGG ACACCTTCC  
8581 CCTCGATCCC CGACGACTGG ACGTTTATAA ACCTAGGCCT CAACAACCGC CCCTTCTTCT  
8641 TCGGCTCGA TGTTAAGAAC TTTACCTTGA ACGCCAACCA AAAGGTTCCC CCCTTAATCG  
8701 TCTATGTCCC CAACCGGCC TATACCGCGC TGAGCAACGT GTCCACCTC GATCGTCA  
8761 ACACCATGTC TCAGCGAAC GACATCATCG GCAACGGATG GAACTCAGCC ACGCAGGGAA  
8821 ACGGCACGCT GGATTGGAG TGGCCCACTT GCGTCGCGCTG CGCGGTTATC AGCAGGAGCT  
8881 TAGATCGGTT GGGCAGGCAG ACGCCAGCCG CGTGAAGAC TTGCTTTGAG AGGTATTGCT  
8941 GGAATGGCAC AGTGAACACTA AAAGATAACAG GGGTTTACAT GCCTGAGTTC AAGATTGCGG  
9001 ATGCGCATGC CCTGGACTCG GGTGCTGTG CTATCGAAA GATGGTGAAT GTCTGGTC  
9061 CGGTTGTGGT GGGAGTTGTG CGGGCTACTT TGTTGTTGTA GGGGTAGGGG AGACGTGATG  
9121 ATATTCCAGT CTGATGAAAGT TGAGACTGGA CTGGAGATCG CCAAGGATGC GGAGGGAAAG  
9181 GAATGCGTGG TGTTAATGTC ATGATGGATG AAGGGTCATG GATCATGGAA CGACGGGGCG

## FIG. 8 continued

9241 GGGATATTGG ATGATGGATA TACCACACTG CATGCATGCT CTATTGATAA TATGCTTG  
9301 CATTTACGTT TAACAATCAA TTGCTCCATC CTGATGTTCT ATCTTCGAC ACTGGATTGA  
9361 TACTACTCCT GTTGCTTCCC TCTTGAAGTT GGAAGGACTT GAGGTTGGAA GGACTTGAGG  
9421 TTGTTTGTTC TGAGGGAGGT TATCGAAGTA TCATCTGTGC TGATGCCGAT CGATAGACTG  
9481 CCCTCTTCTT CGAGGCAACG AACGGTCGGA TGAGCCTCTA ATCATGATGC TCAGTGCCAC  
9541 AAAAAGGCTC CAGCACAGCT GCCCACACCT TTTTGCCTC GTGCCTCCTT CCTTTTTTC  
9601 CCCCCCTTTC TTCCCTTCCA TCTCATCCCG TACCAGAGTG CCCACCGGGT ATATATATTA  
9661 CCTCCTTGGC CGTTCTCCTT TGACCAATAA ATCGCTTGGT CGAGTGGCGT AACCGTTTAC  
9721 CGTCTACACT TATCACTCAA ACCAAACCAA ACCATCGAAG AAGTTACCTA TCGGTTCGAG  
9781 GGAACGGTGA TGTTCTTACG TTCAAGTTAA CCCAAAGAGC GTTCCACATC GTTGAACCGT  
9841 CTCCCTCCAGT TCTTGGATCT GTTTAACTTC CGCAGCGACT GAAGAAGTAA TCACTTTTT  
9901 TTTTTTGTT TCCAAAAAAA AAAAAAAA TTAC

# FIG. 9

**Construction of the components of the sequence diversification cross: Parent (variant 1) and Parent (variant 2).** For convenience, plasmid sequences are shown as linear. The cross hatched region in the chromosome is dispensable. Stippled sequence in the plasmid indicate the multiple cloning site for inserting foreign DNA. Crossovers in region 1 and region 2 insert the foreign sequence to be diversified into chromosome 1 of *Neurospora crassa* adjacent to the recombination hotspot *cog*. Parent (variant 2) containing a version of the foreign sequence with multiple differences from that in parent (variant 1) and parent (variant 2) are crossed and conversion events (stippled arrow) initiated (X) at *cog* permit recombination. *rec-2* on linkage group V permits *cog* to be active. For simplicity, genes not directly related to the diversification are omitted. See text for further details.



## FIG. 10

Construction of parent (variant 1) and parent (variant 2) enabling selection of progeny that have experienced conversion in the foreign DNA. Complementing pairs of *his-3* alleles are used to obtain parent (variant 1) and a different pair of complementing *his-3* alleles are used to obtain parent (variant 2) as explained in the text. Parent (variant 1) and parent (variant 2) are crossed and his<sup>t</sup> recombinants are selected. These must all have experienced conversion events affecting the foreign DNA since the events begin at *cog<sup>L</sup>*. The *his-3* alleles in parent (variant 1) and parent (variant 2) are non complementing to ensure that selection yields recombinants and not aneuploid progeny having two copies of all or part of linkage group 1.

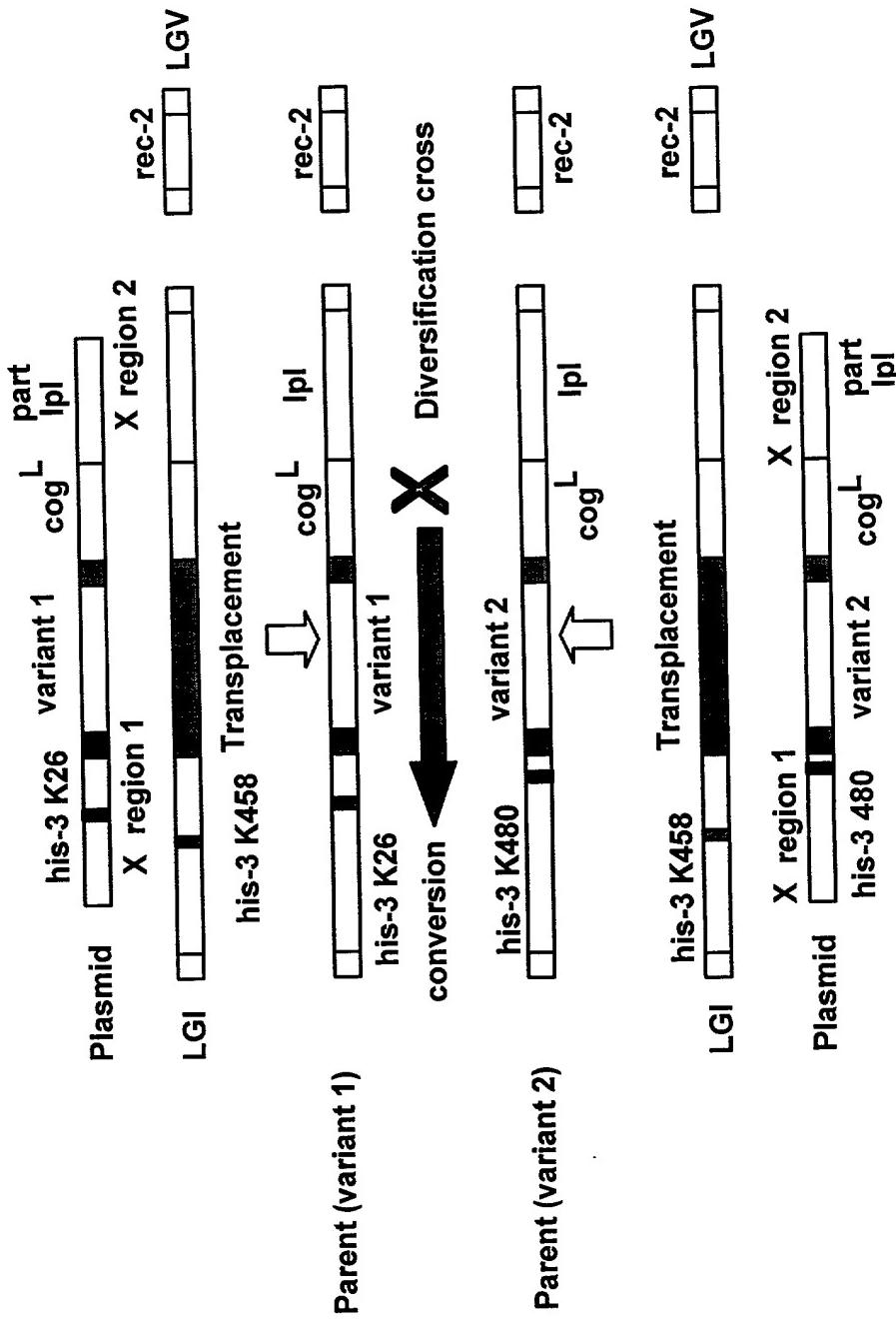


FIG. 11

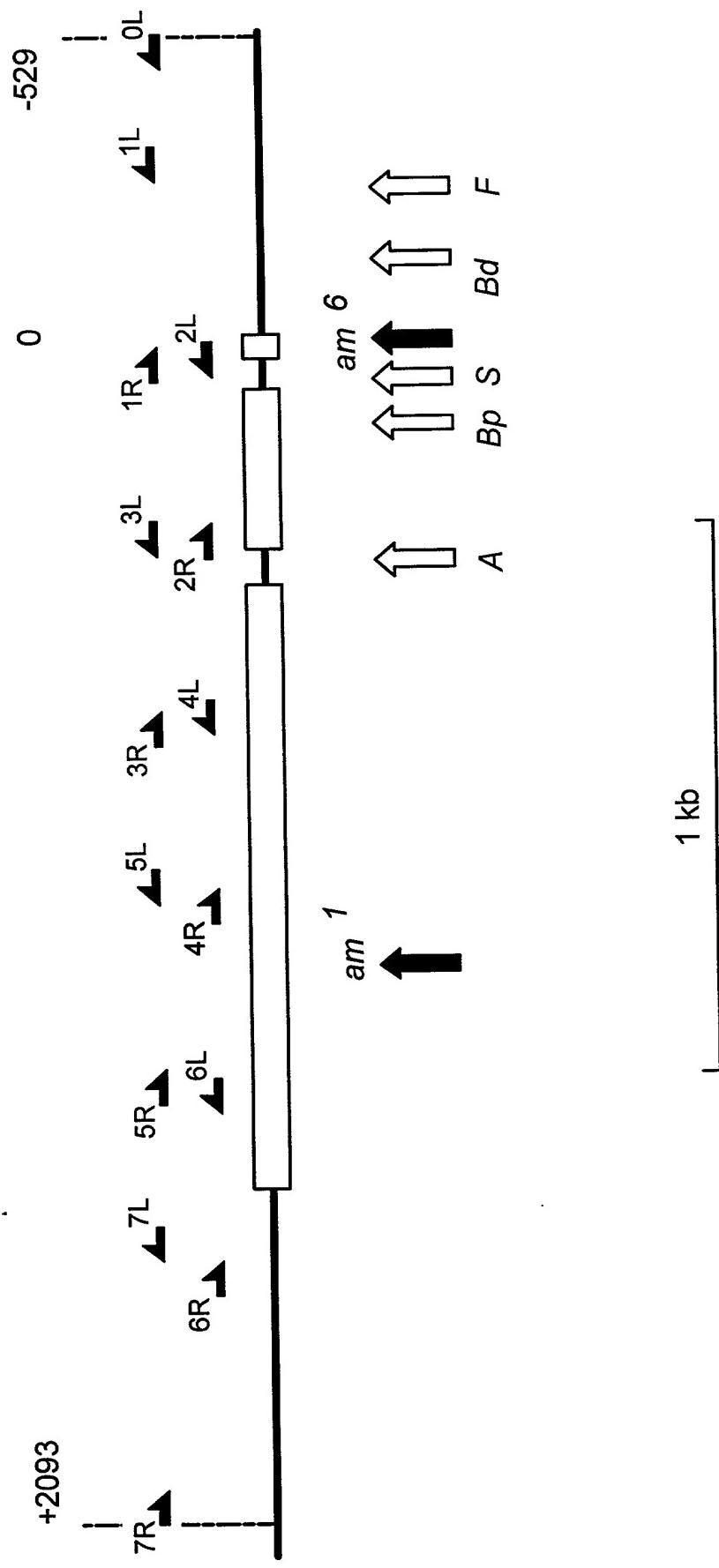


FIG. 12

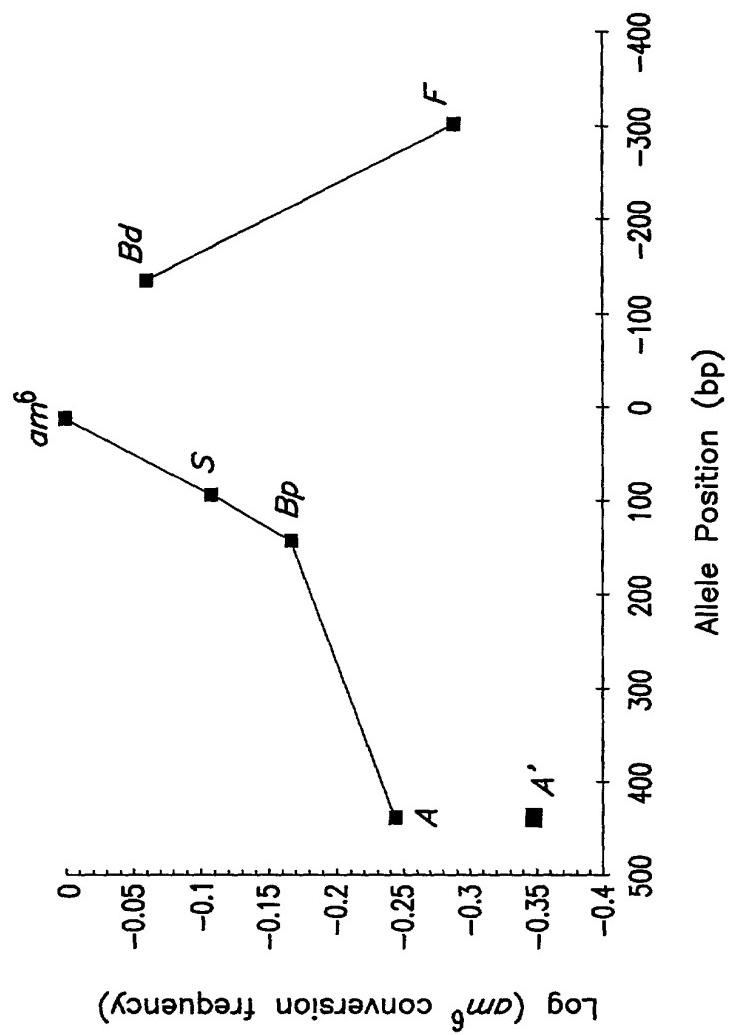
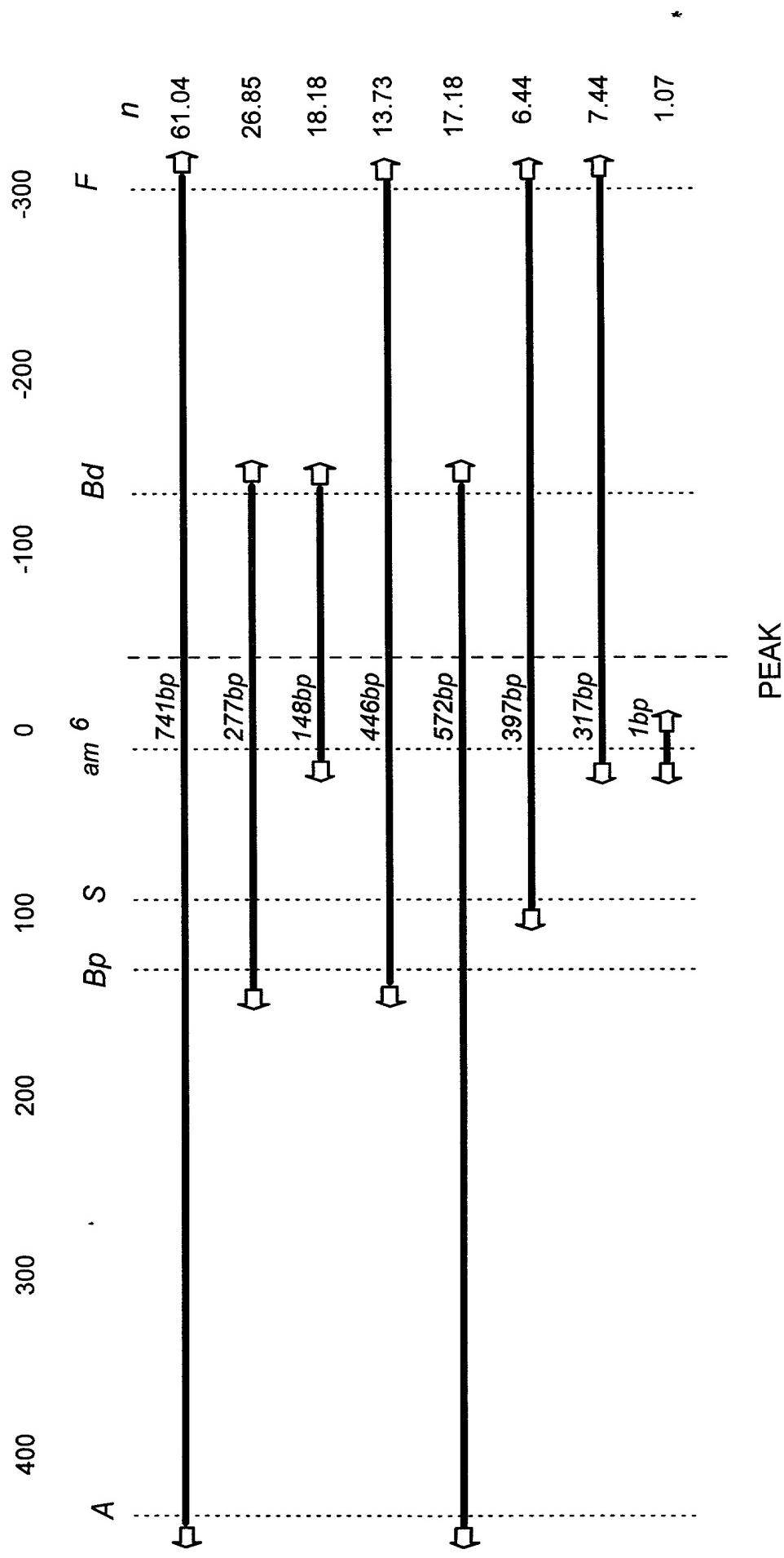


FIG. 13



**FIG. 14**

